

## WEST Search History

DATE: Monday, January 22, 2007

Hide?	<u>Set</u> <u>Name</u>	<u>Query</u>	<u>Hit</u> <u>Count</u>
		<i>DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI; PLUR=YES; OP=OR</i>	
<input type="checkbox"/>	L5	(OPAL1 OR (OPAL ADJ 1)) and polynucleotide	4
		<i>DB=USPT; PLUR=YES; OP=OR</i>	
<input type="checkbox"/>	L4	6979557[PN]	1
		<i>DB=DWPI; PLUR=YES; OP=OR</i>	
<input type="checkbox"/>	L2	(Tang AND Liu AND Asundi AND Xu AND Wehrman AND Ren)[IN]	6
<input type="checkbox"/>	L1	(Yue AND He AND Yao AND Bandman AND Burford AND Tang AND Xu)[IN]	3

END OF SEARCH HISTORY



A service of the National Library of Medicine  
and the National Institutes of Health

My NCBI

[\[Sign In\]](#) [\[Register\]](#)

All Databases

PubMed

Nucleotide

Protein

Genome

Structure

OMIM

PMC

Journals

Books

Search  for

[Limits](#)

[Preview/Index](#)

**[History](#)**

[Clipboard](#)

[Details](#)

About Entrez

Text Version

Entrez PubMed

Overview

Help | FAQ

Tutorials

New/Noteworthy

E-Utilities

PubMed Services

Journals Database

MeSH Database

Single Citation Matcher

Batch Citation Matcher

Clinical Queries

Special Queries

LinkOut

My NCBI

## Search

## Most Recent Queries

## Time Result

[#7](#) Search **IL-10 receptor alpha and leukemia**

10:32:11 [4](#)

[#6](#) Search **G2 and leukemia**

10:28:08 [1069](#)

[#5](#) Search **outcome predictor in acute leukemia**

09:50:24 [149](#)

[#1](#) Search **OPAL1**

09:50:16 [1](#)

[#3](#) Search **Willman CL and leukemia**

09:48:22 [86](#)

[#2](#) Search **G0 and ALL**

09:43:35 [6338](#)

## Related Resources

Order Documents

NLM Mobile

NLM Catalog

NLM Gateway

TOXNET

Consumer Health

Clinical Alerts

ClinicalTrials.gov

PubMed Central

[Write to the Help Desk](#)

[NCBI](#) | [NLM](#) | [NIH](#)

[Department of Health & Human Services](#)

[Privacy Statement](#) | [Freedom of Information Act](#) | [Disclaimer](#)

Jan 16 2007 05:58:20

http://es/ScoreAccessWeb/GetItem.action?AppId=10729895&seqId=1048742&ItemName... 1/22/2007

12: geneseqn2004as:\*  
 13: geneseqn2004bs:\*  
 14: geneseqn2005s:\*  
 15: geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1909	100.0	1080	12	ADP48574	Adp48574 Human OPA
2	1909	100.0	2976	6	AAI72319	Aai72319 ISIGP-2 c
3	1909	100.0	4122	12	ADP48589	Adp48589 Full leng
4	1901	99.6	2703	10	ADA53255	Ada53255 Human cod
5	1868	97.9	1140	12	ADP48576	Adp48576 Human OPA
6	1863	97.6	1508	5	AAS44951	Aas44951 cDNA enco
7	1863	97.6	3095	5	ADL61901	Adl61901 Human ova
8	1853	97.1	4150	4	AAD07816	Aad07816 Human sec
9	1846	96.7	4054	3	AAC59910	Aac59910 Human sec
10	1805	94.6	1715	5	AAS45139	Aas45139 cDNA enco
11	1642.5	86.0	894	12	ADP28830	Adp28830 Human sec
12	1353	70.9	4745	4	AAK65958	Aak65958 Human imm
13	1349	70.7	4744	4	AAK65959	Aak65959 Human imm
14	1294	67.8	69081	11	ACN44642	Acn44642 Human gen
c 15	1178.5	61.7	832	4	AAD07857	Aad07857 Human sec
16	935	49.0	1006	5	ABA21206	Aba21206 Human ner
c 17	918	48.1	505	4	AAI17727	Aai17727 Probe #76
c 18	918	48.1	505	4	ABA62668	Aba62668 Human foe
c 19	918	48.1	505	4	AAI42669	Aai42669 Probe #11
c 20	918	48.1	505	4	ABA29977	Aba29977 Probe #84
c 21	918	48.1	505	4	AAK36874	Aak36874 Human bon
c 22	918	48.1	505	4	AAK11046	Aak11046 Human bra
c 23	918	48.1	505	4	ABS36548	Abs36548 Human liv
c 24	918	48.1	505	6	ABS10887	Abs10887 Human gen
c 25	753	39.4	84973	12	ADH69807_6	Continuation (7 of
26	751	39.3	439	5	ADL40673	Adl40673 Human ova
27	747	39.1	577	8	ABZ36387	Abz36387 Human sec
28	719.5	37.7	605	4	AAD07856	Aad07856 Human sec
29	715	37.5	550	4	AAH98928	Aah98928 Human EST
30	650	34.0	435	13	ADQ53832	Adq53832 Novel can
31	642	33.6	603	13	ADU02259	Adu02259 Novel hum
c 32	640	33.5	352	4	AAI26927	Aai26927 Probe #16
c 33	640	33.5	352	4	ABA75183	Aba75183 Human foe
c 34	640	33.5	352	4	AAI55731	Aai55731 Probe #24
c 35	640	33.5	352	4	ABA39851	Aba39851 Probe #18
c 36	640	33.5	352	4	AAK49818	Aak49818 Human bon
c 37	640	33.5	352	4	AAK23717	Aak23717 Human bra
c 38	640	33.5	352	4	ABS49458	Abs49458 Human liv
c 39	640	33.5	352	6	ABS23320	Abs23320 Human gen
40	631	33.1	382	5	ADI75437	Adi75437 Human ova
41	631	33.1	382	5	ADI69090	Adi69090 Human ova
42	579	30.3	462	9	ACH22664	Ach22664 Human adu
43	476.5	25.0	1183	14	ADY18021	Ady18021 DNA encod
44	476.5	25.0	1199	3	AAC77716	Aac77716 Human can
45	474.5	24.9	1024	3	AAA60194	Aaa60194 Hydrophob

#### ALIGNMENTS

##### RESULT 1

ADP48574

ID ADP48574 standard; DNA; 1080 BP.

XX

AC ADP48574;

XX

DT 09-SEP-2004 (first entry)

XX

DE Human OPAL1/G0 DNA splice form incorporating exon 1 SeqID 1.

XX

KW human; gene; ds; OPAL1; G0; childhood leukaemia; G protein beta 2; G1;

KW interleukin-10 receptor alpha; G2; IL-10; FYN-binding protein; PBK1;

KW cytostatic; acute lymphoblastic leukaemia; ALL; acute myeloid leukaemia;



KW AML.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1029  
FT /\*tag= a  
FT /product= "OPAL1/G0 protein (incorporating exon 1)"  
XX  
PN WO2004053074-A2.  
XX  
PD 24-JUN-2004.  
XX  
PF 05-DEC-2003; 2003WO-US038738.  
XX  
PR 06-DEC-2002; 2002US-0432064P.  
PR 06-DEC-2002; 2002US-0432077P.  
PR 06-DEC-2002; 2002US-0432078P.  
PR 14-OCT-2003; 2003US-0510904P.  
PR 14-OCT-2003; 2003US-0510968P.  
PR 05-DEC-2003; 2003US-00729895.  
PR 05-DEC-2003; 2003US-0527610P.  
XX  
PA (SCTE-) SCI & TECHNOLOGY CORP @UNM.  
PA (SAND-) SANDIA CORP.  
XX  
PI Willman CL, Helman P, Veroff R, Mosquera-Caro M, Davidson GS;  
PI Martin SB, Atlas SR, Andries E, Kang H, Shuster JJ, Haaland DM;  
PI Potter JW, Wang X, Harvey R;  
XX  
DR WPI; 2004-468846/44.  
DR P-PSDB; ADP48575.  
XX  
PT New isolated Outcome Predictor in Acute Leukemia 1 (OPAL1)  
PT polynucleotides and polypeptides for preventing, treating or diagnosing  
PT leukemia, or for screening or evaluating compounds that may be used for  
PT treating leukemia.  
XX  
PS Claim 1; SEQ ID NO 1; 343pp; English.  
XX  
CC This invention relates to a novel isolated nucleic acid identified as  
CC OPAL1 (also known as G0), and the encoded protein thereof. Specifically,  
CC it refers to a method for predicting the outcome and risk classification  
CC of childhood leukaemia. The present invention describes a gene profiling  
CC method that can be used to determine the expression level for OPAL1, a  
CC gene that is strongly predictive regarding the outcome of both acute  
CC lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML).  
CC Furthermore, it provides additional genes, namely G protein beta 2  
CC (referred to as G1), interleukin-10 (IL-10) receptor alpha (referred to  
CC as G2), the FYN-binding protein and PBK1 that can also be used, alone or  
CC in combination, in predictive assays for the classification and potential  
CC therapy for a leukaemia patient. Accordingly, these compositions exhibit  
CC cytostatic activities. This polynucleotide sequence is the human OPAL1/G0  
CC splice from incorporating exon 1 of the invention.  
XX  
SQ Sequence 1080 BP; 249 A; 341 C; 292 G; 198 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.64e-104	Length:	1080
Score:	1909.00	Matches:	342
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-729-895A-2 (1-342) x ADP48574 (1-1080)

Qy	1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	1 ATGCCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	60
Qy	21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	120
Qy	41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60

```

Db      121 |||||GAAGTCTGGTGGTCTGGCTGGTGGACCATCATCATCATCTGAGCTGCTGCTGTGT 180
Qy      61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80
Db      181 |||||TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240
Qy      81 |||||LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100
Db      241 |||||CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTG 300
Qy      101 |||||ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120
Db      301 |||||CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA 360
Qy      121 |||||ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140
Db      361 |||||CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 420
Qy      141 |||||GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160
Db      421 |||||GGTGGCAGTCCCCGGGCATCGATCCACAGGGGATCCAGGGGGACAGAGCAGCCCC 480
Qy      161 |||||LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180
Db      481 |||||TTGTCTGAGCCAGCAGAGAAGCAGCACAAAGCCCCAAGCATCGCTGACCCTGATCCCTCT 540
Qy      181 |||||AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200
Db      541 |||||GACCTACCAGTTGACCGAGCAGCCACAAAGCCCCAGGATGGAGCCAGTGGCTCTGTG 600
Qy      201 |||||AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220
Db      601 |||||GCTGGCCTGGGGGAGCTGGACCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
Qy      221 |||||GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240
Db      661 |||||GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
Qy      241 |||||GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260
Db      721 |||||GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 780
Qy      261 |||||HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280
Db      781 |||||CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC 840
Qy      281 |||||LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300
Db      841 |||||CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGCCTCTGT 900
Qy      301 |||||GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320
Db      901 |||||CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
Qy      321 |||||CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340
Db      961 |||||TGCTGTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1020
Qy      341 |||||ProSer 342
Db      1021 |||||CCCAGC 1026

```

RESULT 2  
AAI72319

ID AAI72319 standard; cDNA; 2976 BP.

XX

AC AAI72319;

XX

DT 15-APR-2002 (first entry)

XX

DE ISIGP-2 cDNA.

XX

KW Human; intracellular signalling protein; ISIGP; gene; cell proliferation;  
autoimmune; inflammation; gastrointestinal disorder;

KW reproductive disorder; developmental disorder; ss.

XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 479..1507  
FT /\*tag= a  
FT /product= "ISIGP-2"  
XX  
PN WO200194391-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 07-JUN-2001; 2001WO-US018595.  
XX  
PR 08-JUN-2000; 2000US-0210582P.  
PR 16-JUN-2000; 2000US-0212443P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Yue H, He A, Nguyen DB, Yao MG, Bandman O, Burford N, Tang YT;  
PI Xu Y, Hafalia A, Azimzai Y, Walia NK;  
XX  
DR WPI; 2002-154564/20.  
DR P-PSDB; AAB47872.  
XX  
PT New human intracellular signaling protein and polynucleotides useful for  
PT diagnosing, treating or preventing cell proliferative,  
PT autoimmune/inflammatory, gastrointestinal, reproductive and developmental  
PT disorders.  
XX  
PS Claim 5; Page 102-03; 106pp; English.  
XX  
CC The sequences given in AAI72318-22 encode novel human intracellular  
CC signalling proteins (ISIGP). The polynucleotides and ISIGP proteins may  
CC be used for the diagnosis, treatment or prevention of cell proliferative,  
CC autoimmune/inflammatory, gastrointestinal, reproductive and developmental  
CC disorders. The protein encoded by this sequence has homology to human WW  
CC domain binding protein-1  
XX  
SQ Sequence 2976 BP; 633 A; 817 C; 823 G; 703 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.81e-104	Length:	2976
Score:	1909.00	Matches:	342
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-729-895A-2 (1-342) x AAI72319 (1-2976)

Qy	1	MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	479	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	538
Qy	21	SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	539	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	598
Qy	41	GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60
Db	599	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGT	658
Qy	61	CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn	80
Db	659	TGCCACCACCGCCGAGCCAAGCACCAGCCTTCAGGCCAGCAGCGCAACATGAAATCAAC	718
Qy	81	LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu	100
Db	719	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTG	778
Qy	101	ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro	120
Db	779	CCAAACTATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCCCCA	838
Qy	121	ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla	140

```

Db      839  |||||||
      839  CCATACAGTGCCTTCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 898
Qy      141  GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160
      141  |||||||
Db      899  GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC 958
Qy      161  LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180
      161  |||||||
Db      959  TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 1018
Qy      181  AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200
      181  |||||||
Db      1019 GACCTACCAGTTGACCGAGCAGCCACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 1078
Qy      201  AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220
      201  |||||||
Db      1079 GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTTGACAAAGATGCAGAATGTAGGGAG 1138
Qy      221  GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240
      221  |||||||
Db      1139 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 1198
Qy      241  GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260
      241  |||||||
Db      1199 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 1258
Qy      261  HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280
      261  |||||||
Db      1259 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC 1318
Qy      281  LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300
      281  |||||||
Db      1319 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1378
Qy      301  GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320
      301  |||||||
Db      1379 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1438
Qy      321  CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340
      321  |||||||
Db      1439 TGCTGTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1498
Qy      341  ProSer 342
      341  |||||
Db      1499 CCCAGC 1504

```

# RESULT 3

ADP48589

ID ADP48589 standard; cDNA; 4122 BP.

XX

AC ADP48589;

XX

DT 09-SEP-2004 (first entry)

XX

DE Full length human OPAL1 cDNA SeqID 16.

XX

KW human; ss; OPAL1; G0; childhood leukaemia; G protein beta 2; G1;  
 KW interleukin-10 receptor alpha; G2; IL-10; FYN-binding protein; PBK1;  
 KW cytostatic; acute lymphoblastic leukaemia; ALL; acute myeloid leukaemia;  
 KW AML.

XX

OS Homo sapiens.

XX

PN WO2004053074-A2.

XX

PD 24-JUN-2004.

XX

PF 05-DEC-2003; 2003WO-US038738.

XX

PR 06-DEC-2002; 2002US-0432064P.

PR 06-DEC-2002; 2002US-0432077P.

PR 06-DEC-2002; 2002US-0432078P.

PR 14-OCT-2003; 2003US-0510904P.

PR 14-OCT-2003; 2003US-0510968P.

PR 05-DEC-2003; 2003US-00729895.

PR 05-DEC-2003; 2003US-0527610P.  
XX  
PA (SCTE-) SCI & TECHNOLOGY CORP @UNM.  
PA (SAND-) SANDIA CORP.  
XX  
PI Willman CL, Helman P, Veroff R, Mosquera-Caro M, Davidson GS;  
PI Martin SB, Atlas SR, Andries E, Kang H, Shuster JJ, Haaland DM;  
PI Potter JW, Wang X, Harvey R;  
XX

DR WPI; 2004-468846/44.  
XX

PT New isolated Outcome Predictor in Acute Leukemia 1 (OPAL1)  
PT polynucleotides and polypeptides for preventing, treating or diagnosing  
PT leukemia, or for screening or evaluating compounds that may be used for  
PT treating leukemia.  
XX

PS Disclosure; SEQ ID NO 16; 343pp; English.  
XX

CC This invention relates to a novel isolated nucleic acid identified as  
CC OPAL1 (also known as G0), and the encoded protein thereof. Specifically,  
CC it refers to a method for predicting the outcome and risk classification  
CC of childhood leukaemia. The present invention describes a gene profiling  
CC method that can be used to determine the expression level for OPAL1, a  
CC gene that is strongly predictive regarding the outcome of both acute  
CC lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML).  
CC Furthermore, it provides additional genes namely G protein beta 2  
CC (referred to as G1), interleukin-10 (IL-10) receptor alpha (referred to  
CC as G2), the FYN-binding protein and PBK1 that can also be used, alone or  
CC in combination, in predictive assays for the classification and potential  
CC therapy for a leukaemia patient. Accordingly, these compositions exhibit  
CC cytostatic activities. This polynucleotide sequence is the full length  
CC human OPAL1 cDNA of the invention.  
XX

SQ Sequence 4122 BP; 964 A; 1066 C; 1075 G; 1017 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	6.81e-104	Length:	4122
Score:	1909.00	Matches:	342
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-729-895A-2 (1-342) x ADP48589 (1-4122)

```
QY      1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
      |||
DB      145 ATGCCCTTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGTACCAACAATCAA 204

QY      21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40
      |||
DB      205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264

QY      41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60
      |||
DB      265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 324

QY      61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80
      |||
DB      325 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384

QY      81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100
      |||
DB      385 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCTGAGTTTGTG 444

QY      101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120
      |||
DB      445 CCAAACTATTTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA 504

QY      121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140
      |||
DB      505 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 564

QY      141 GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160
      |||
DB      565 GGTGGCAGTCCCCCGGCATCGATCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC 624
```

Qy 161 LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180  
 |||||  
 Db 625 TTGTCTGAGCCAGCAGAAGCAGCACAAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 684  
 Qy 181 AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200  
 |||||  
 Db 685 GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCGAGGATGGAGCCCACTGGCTCTGTG 744  
 Qy 201 AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220  
 |||||  
 Db 745 GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 804  
 Qy 221 GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240  
 |||||  
 Db 805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864  
 Qy 241 GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260  
 |||||  
 Db 865 GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924  
 Qy 261 HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280  
 |||||  
 Db 925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC 984  
 Qy 281 LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300  
 |||||  
 Db 985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044  
 Qy 301 GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320  
 |||||  
 Db 1045 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1104  
 Qy 321 CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340  
 |||||  
 Db 1105 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1164  
 Qy 341 ProSer 342  
 |||||  
 Db 1165 CCCAGC 1170

RESULT 4

ADA53255

ID ADA53255 standard; cDNA; 2703 BP.

XX

AC ADA53255;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human coding sequence, SEQ ID 823.

XX

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

KW Gene Therapy; human; secretory protein; membrane proteins; cancer;

KW inflammatory disease; osteoporosis; neurological disease; gene; ss.

XX

OS Homo sapiens.

XX

PN EP1293569-A2.

XX

PD 19-MAR-2003.

XX

PF 21-MAR-2002; 2002EP-00006586.

XX

PR 14-SEP-2001; 2001JP-00328381.

PR 24-JAN-2002; 2002US-0350435P.

XX

PA (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX

DR WPI; 2003-395539/38.

DR P-PSDB; ADA54894.

XX

PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX  
PS Claim 1; SEQ ID NO 823; 205pp; English.  
XX  
CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
XX  
SQ Sequence 2703 BP; 599 A; 742 C; 713 G; 649 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.3e-103	Length:	2703
Score:	1901.00	Matches:	340
Percent Similarity:	99.7%	Conservative:	1
Best Local Similarity:	99.4%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	10	Gaps:	0

US-10-729-895A-2 (1-342) x ADA53255 (1-2703)

Qy	1	MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	145	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	204
Qy	21	SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	205	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	264
Qy	41	GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60
Db	265	GAACCTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	324
Qy	61	CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn	80
Db	325	TGCCACCACCGCCGAGCCAAGCACGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	384
Qy	81	LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu	100
Db	385	CTGATCGCTTACCGAGAAGCCACAACTACTCAGCGCTGCCATTTTATTTTCAAGTTTTTG	444
Qy	101	ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro	120
Db	445	CCAAACTATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA	504
Qy	121	ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla	140
Db	505	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	564
Qy	141	GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro	160
Db	565	GGTGGCAGTCCCCGGGCATCGATCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC	624
Qy	161	LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer	180
Db	625	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	684
Qy	181	AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal	200
Db	685	GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	744
Qy	201	AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu	220
Db	745	GCTGGCCTGGGGGAGCTGGACCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	804
Qy	221	GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro	240
Db	805	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	864
Qy	241	GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly	260
Db	865	GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG	924

Qy 261 HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280  
 |||  
 Db 925 CACCATGACGATGACCTCAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC 984

Qy 281 LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300  
 |||  
 Db 985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy 301 GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320  
 |||  
 Db 1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCTGGGCACCCGCACCTGCCACGGCCGCCCTCA 1104

Qy 321 CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340  
 |||  
 Db 1105 TGCTGTGCTGAACACCATCAACGAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC 1164

Qy 341 ProSer 342  
 |||  
 Db 1165 CCCAGC 1170

RESULT 5

ADP48576

ID ADP48576 standard; DNA; 1140 BP.

XX

AC ADP48576;

XX

DT 09-SEP-2004 (first entry)

XX

DE Human OPAL1/G0 DNA splice form incorporating exon 1a SeqID 3.

XX

KW human; gene; ds; OPAL1; G0; childhood leukaemia; G protein beta 2; G1;

KW interleukin-10 receptor alpha; G2; IL-10; FYN-binding protein; PBK1;

KW cytostatic; acute lymphoblastic leukaemia; ALL; acute myeloid leukaemia;

KW AML.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1092

FT /\*tag= a

FT /product= "OPAL1/G0 protein (incorporating exon 1a)"

XX

PN WO2004053074-A2.

XX

PD 24-JUN-2004.

XX

PF 05-DEC-2003; 2003WO-US038738.

XX

PR 06-DEC-2002; 2002US-0432064P.

PR 06-DEC-2002; 2002US-0432077P.

PR 06-DEC-2002; 2002US-0432078P.

PR 14-OCT-2003; 2003US-0510904P.

PR 14-OCT-2003; 2003US-0510968P.

PR 05-DEC-2003; 2003US-00729895.

PR 05-DEC-2003; 2003US-0527610P.

XX

PA (SCTE-) SCI & TECHNOLOGY CORP @UNM.

PA (SAND-) SANDIA CORP.

XX

PI Willman CL, Helman P, Veroff R, Mosquera-Caro M, Davidson GS;

PI Martin SB, Atlas SR, Andries E, Kang H, Shuster JJ, Haaland DM;

PI Potter JW, Wang X, Harvey R;

XX

DR WPI; 2004-468846/44.

DR P-PSDB; ADP48577.

XX

PT New isolated Outcome Predictor in Acute Leukemia 1 (OPAL1)

PT polynucleotides and polypeptides for preventing, treating or diagnosing

PT leukemia, or for screening or evaluating compounds that may be used for

PT treating leukemia.

XX

PS Claim 1; SEQ ID NO 3; 343pp; English.

XX

CC This invention relates to a novel isolated nucleic acid identified as

CC OPAL1 (also known as G0), and the encoded protein thereof. Specifically,

CC it refers to a method for predicting the outcome and risk classification



CC of childhood leukaemia. The present invention describes a gene profiling  
 CC method that can be used to determine the expression level for OPAL1, a  
 CC gene that is strongly predictive regarding the outcome of both acute  
 CC lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML).  
 CC Furthermore, it provides additional genes namely G protein beta 2  
 CC (referred to as G1), interleukin-10 (IL-10) receptor alpha (referred to  
 CC as G2), the FYN-binding protein and PBK1 that can also be used, alone or  
 CC in combination, in predictive assays for the classification and potential  
 CC therapy for a leukaemia patient. Accordingly, these compositions exhibit  
 CC cytostatic activities. This polynucleotide sequence is the human OPAL1/GO  
 CC splice from incorporating exon 1a of the invention.

XX

SQ Sequence 1140 BP; 258 A; 367 C; 314 G; 201 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.7e-102	Length:	1140
Score:	1868.00	Matches:	334
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	97.9%	Indels:	0
DB:	12	Gaps:	0

US-10-729-895A-2 (1-342) x ADP48576 (1-1140)

Qy	9	GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis	28
Db	88	CAGGATAAGGAAGCCTGTGTGGGTACCAACATCAAAGCTACATCTGTGACACAGGACAC	147
Qy	29	CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal	48
Db	148	TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAAGCTGGTGGTTCTGGCTGGTG	207
Qy	49	TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis	68
Db	208	TGGACCATCATCATCATCCTGAGTGCTGCTGTTGCCACACCGCCGAGCCAAGCAC	267
Qy	69	ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis	88
Db	268	CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC	327
Qy	89	AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr	108
Db	328	AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAAACTATTACTACCTCCTTAT	387
Qy	109	GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln	128
Db	388	GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCTTCCAGCTACAG	447
Qy	129	GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp	148
Db	448	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT	507
Qy	149	ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer	168
Db	508	CCCACCAGGGGATCCCAGGGGGCACAGACAGCCCTTGTCTGAGCCAGCAGAAGCAGC	567
Qy	169	ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla	188
Db	568	ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACAGTTGACCGAGCAGCC	627
Qy	189	ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro	208
Db	628	ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCTGGGGAGCTGGACCCG	687
Qy	209	GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer	228
Db	688	GGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT	747
Qy	229	GluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGly	248
Db	748	GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTACAGGT	807
Qy	249	AspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGlu	268
Db	808	GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG	867

Qy 269 PheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHis 288  
 |||  
 Db 868 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT 927  
 Qy 289 ValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArg 308  
 |||  
 Db 928 GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCTCCTCTGAGGAGCAGGCTCGA 987  
 Qy 309 GluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsn 328  
 |||  
 Db 988 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCGCATGCCTGCTGCTGAACACCATCAAC 1047  
 Qy 329 GluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342  
 |||  
 Db 1048 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCCAGG 1089

RESULT 6

AAS44951

ID AAS44951 standard; cDNA; 1508 BP.

XX

AC AAS44951;

XX

DT 18-DEC-2001 (first entry)

XX

DE cDNA encoding novel human secretory protein, Seq ID No 32.

XX

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
 KW fertility; analgesic; pain; antigen; ss.

XX

OS Homo sapiens.

XX

PN WO200166689-A2.

XX

PD 13-SEP-2001.

XX

PF 05-MAR-2001; 2001WO-US004942.

XX

PR 07-MAR-2000; 2000US-00519705.

PR 19-MAY-2000; 2000US-00574454.

PR 17-JUN-2000; 2000US-00596193.

PR 14-JUL-2000; 2000US-00616847.

PR 19-SEP-2000; 2000US-00665363.

PR 20-OCT-2000; 2000US-00693267.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX

DR WPI; 2001-589934/66.

DR P-PSDB; AAU28051.

XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues, for diagnosis and treatment of  
 PT cancer, neurological, inflammatory, and autoimmune disorders.

XX

PS Claim 1; SEQ ID NO 32; 107pp; English.

XX

CC The invention relates to novel isolated human secreted polypeptides (I)  
 CC and polynucleotides (II). (I) and (II) are useful for treating  
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
 CC involved in increasing haematopoiesis, stem cell survival, bone growth  
 CC and remodeling. (I), (II) and modulators of (II) are useful for  
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
 CC creating transgenic animals useful for studying the in vivo activities of  
 CC the polypeptide as well as for studying modulators of the polypeptides.  
 CC (I) induces the proliferation of neural cells and regeneration of nerve  
 CC and brain tissue and is useful for the treatment of central and

CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,  
 CC or periodontal disease. Furthermore, (I) is also useful for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues, various immune deficiencies and  
 CC disorders including severe combined immunodeficiency (SCID), bacterial or  
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,  
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
 CC reactions and conditions, such as asthma or other respiratory problems.  
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,  
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
 CC analgesic effects or other pain reducing effects, immunoglobulin like  
 CC activity and can act as an antigen in a vaccine composition to raise an  
 CC immune response. AAS44920-AAS45295 represent novel human secreted protein  
 CC coding sequences of the invention

XX

SQ Sequence 1508 BP; 345 A; 447 C; 423 G; 293 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	1.25e-101	Length:	1508
Score:	1863.00	Matches:	333
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	97.6%	Indels:	0
DB:	5	Gaps:	0

US-10-729-895A-2 (1-342) x AAS44951 (1-1508)

Qy	9	GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis	28
Db	174	CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC	233
Qy	29	CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal	48
Db	234	TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAAGTCTGGTGGTTCTGGCTGGTG	293
Qy	49	TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis	68
Db	294	TGGACCATCATCATCATCCTGAGTGCTGCTGTTTGGCCACCACGCCGAGCCAAGCAC	353
Qy	69	ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis	88
Db	354	CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC	413
Qy	89	AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr	108
Db	414	AATTACTCAGCGCTGCCATTTTATTTAGGTTTTGCCAACTATTACTACCTCCTTAT	473
Qy	109	GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln	128
Db	474	GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCTTCCAGCTACAG	533
Qy	129	GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp	148
Db	534	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT	593
Qy	149	ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer	168
Db	594	CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC	653
Qy	169	ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla	188
Db	654	ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAAGTTGACCGAGCAGCC	713
Qy	189	ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro	208
Db	714	ACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG	773
Qy	209	GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer	228

```

Db      774  |||||GGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 833
Qy      229  GluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGly 248
Db      834  |||||GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 893
Qy      249  AspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGlu 268
Db      894  |||||GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 953
Qy      269  PheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHis 288
Db      954  |||||TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGTGCCAT 1013
Qy      289  ValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArg 308
Db      1014 |||||GTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGTCAGCCCTCTGAGGAGCAGGCTCGA 1073
Qy      309  GluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsn 328
Db      1074 |||||GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1133
Qy      329  GluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342
Db      1134 |||||GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCCAGC 1175

```

RESULT 7

ADL61901

ID ADL61901 standard; DNA; 3095 BP.

XX

AC ADL61901;

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006



12: gb\_htg:\*  
13: gb\_in:\*  
14: gb\_om:\*  
15: gb\_ba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1909	100.0	2976	2	AX337966	AX337966 Sequence
2	1901	99.6	2703	2	AX714139	AX714139 Sequence
3	1901	99.6	2703	5	AK056285	AK056285 Homo sapi
4	1863	97.6	1002	2	CQ720290	CQ720290 Sequence
5	1863	97.6	3095	2	CQ413042	CQ413042 Sequence
6	1846	96.7	4054	2	BD275013	BD275013 50 Human
7	1618	84.8	2241	6	BC026369	BC026369 Mus muscu
8	1618	84.8	3921	6	BC058949	BC058949 Mus muscu
9	1353	70.9	131753	5	AL358790	AL358790 Human DNA
10	1353	70.9	173758	5	AC009144	AC009144 Homo sapi
11	1294	67.8	62035	5	AC078937	AC078937 Homo sapi
12	1293	67.7	228263	12	AC079930	AC079930 Homo sapi
13	1275	66.8	170261	5	AC147974	AC147974 Pan trogl
14	1172	61.4	2217	5	AK000374	AK000374 Homo sapi
15	1159	60.7	275121	12	AC099420	AC099420 Rattus no
c 16	1149.5	60.2	206019	6	AC161865	AC161865 Mus muscu
17	1149.5	60.2	216311	6	AC156982	AC156982 Mus muscu
18	1085.5	56.9	196003	12	AC079419	AC079419 Mus muscu
c 19	1025.5	53.7	74874	12	AC168985	AC168985 Bos tauru
20	1017.5	53.3	150528	12	AC149242	AC149242 Macaca mu
21	1017.5	53.3	166044	12	AC150715	AC150715 Macaca mu
c 22	1017.5	53.3	176436	12	AC148535	AC148535 Macaca mu
c 23	994.5	52.1	239579	12	AC095597	AC095597 Rattus no
c 24	994.5	52.1	247659	12	AC095548	AC095548 Rattus no
25	982.5	51.5	247169	12	AC123007	AC123007 Rattus no
c 26	918	48.1	505	2	CQ071860	CQ071860 Sequence
c 27	918	48.1	505	2	CQ102496	CQ102496 Sequence
c 28	918	48.1	505	2	CQ141409	CQ141409 Sequence
c 29	918	48.1	505	2	CQ177047	CQ177047 Sequence
c 30	918	48.1	505	2	CQ224699	CQ224699 Sequence
c 31	918	48.1	505	2	CQ262712	CQ262712 Sequence
c 32	918	48.1	505	2	CQ299773	CQ299773 Sequence
c 33	918	48.1	505	2	CQ336943	CQ336943 Sequence
34	819.5	42.9	1479	11	BC095609	BC095609 Danio rer
c 35	768.5	40.3	84108	12	AC133385	AC133385 Rattus no
36	768.5	40.3	166247	12	AC135876	AC135876 Rattus no
c 37	768.5	40.3	218780	12	AC115552	AC115552 Rattus no
38	753	39.4	28657	12	AC153482	AC153482 Homo sapi
c 39	753	39.4	35944	5	HSTCRB1H	AF009664 Homo sapi
40	753	39.4	67826	5	AC099395	AC099395 Homo sapi
c 41	753	39.4	161425	5	AC104597	AC104597 Homo sapi
c 42	753	39.4	232650	5	U66061	U66061 Human germl
43	751	39.3	439	2	CQ407492	CQ407492 Sequence
44	652	34.2	984	2	CQ716136	CQ716136 Sequence
45	648	33.9	169309	6	AL731692	AL731692 Mouse DNA

#### ALIGNMENTS

RESULT 1  
AX337966  
LOCUS AX337966 2976 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 7 from Patent WO0194391.  
ACCESSION AX337966  
VERSION AX337966.1. GI:18128677  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1

AUTHORS Yue,H., He,A., Nguyen,D.B., Yao,M.G., Bandman,O., Burford,N.,  
Tang,Y.T., Xu,Y., Hafalia,A., Azimzai,Y. and Walia,N.K.  
TITLE Intracellular signaling proteins  
JOURNAL Patent: WO 0194391-A 7 13-DEC-2001;  
Incyte Genomics, Inc. (US)  
FEATURES Location/Qualifiers  
source 1. 2976  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 1478005CB1"

# ORIGIN

## Alignment Scores:

Pred. No.:	2.84e-128	Length:	2976
Score:	1909.00	Matches:	342
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-729-895A-2 (1-342) x AX337966 (1-2976)

Qy	1	MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	479	ATGCCTTTCCTTTTGGGTCTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	538
Qy	21	SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	539	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	598
Qy	41	GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60
Db	599	GAAGTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	658
Qy	61	CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn	80
Db	659	TGCCACCACCGCCGAGCCAAAGCACCAGCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	718
Qy	81	LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu	100
Db	719	CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTGTG	778
Qy	101	ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro	120
Db	779	CCAACTATTTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA	838
Qy	121	ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla	140
Db	839	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	898
Qy	141	GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro	160
Db	899	GGTGGCAGTCCCCGGGCATCGATCCACACAGGGGATCCAGGGGGCAGAGCAGCCCC	958
Qy	161	LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer	180
Db	959	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	1018
Qy	181	AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal	200
Db	1019	GACCTACAGTTGACCGAGCAGCCACAAAGCCCGAGGATGGAGCCAGTGGCTCTGTG	1078
Qy	201	AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu	220
Db	1079	GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	1138
Qy	221	GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro	240
Db	1139	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	1198
Qy	241	GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly	260
Db	1199	GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGCAACCGGGGC	1258
Qy	261	HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro	280

```

Db      1259  |||||CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 1318
Qy      281  LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300
Db      1319  |||||CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1378
Qy      301  |||||GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320
Db      1379  |||||CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1438
Qy      321  |||||CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340
Db      1439  |||||TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC 1498
Qy      341  ProSer 342
Db      1499  |||||CCCAGC 1504

```

## RESULT 2

AX714139

LOCUS AX714139 2703 bp DNA linear PAT 15-APR-2003

DEFINITION Sequence 823 from Patent EP1293569.

ACCESSION AX714139

VERSION AX714139.1 GI:29889067

## KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

## REFERENCE

1  
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,  
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,  
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and  
Masuho,Y.

TITLE Full-length cDNAs

JOURNAL Patent: EP 1293569-A 823 19-MAR-2003;  
Helix Research Institute (JP) ; Research Association for  
Biotechnology (JP)

## FEATURES

source Location/Qualifiers  
1..2703  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.:	9.67e-128	Length:	2703
Score:	1901.00	Matches:	340
Percent Similarity:	99.7%	Conservative:	1
Best Local Similarity:	99.4%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	2	Gaps:	0

US-10-729-895A-2 (1-342) x AX714139 (1-2703)

```

Qy      1  MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
Db      145  |||||ATGCCCTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACATCAA 204
Qy      21  SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40
Db      205  |||||AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264
Qy      41  GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60
Db      265  |||||GAACTCTGGTGGTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 324
Qy      61  CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80
Db      325  |||||TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGCAACATGAAATCAAC 384
Qy      81  LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100
Db      385  |||||CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTCAGGTTTGTG 444

```



Qy 101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120  
 |||||  
 Db 445 CCAAACTATTCTACTACCTCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA 504

Qy 121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140  
 |||||  
 Db 505 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 564

Qy 141 GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160  
 |||||  
 Db 565 GGTGGCAGTCCCCGGGCATCGATCCACCAGGGGATCCAGGGGGCACAGAGCAGCCCC 624

Qy 161 LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180  
 |||||  
 Db 625 TTGTCTGAGCCAGCAGAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 684

Qy 181 AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200  
 |||||  
 Db 685 GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 744

Qy 201 AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220  
 |||||  
 Db 745 GCTGGCCTGGGGGAGCTGGACCCGGGGGCTTCTGGACAAAGATGCAGAATGTAGGGAG 804

Qy 221 GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240  
 |||||  
 Db 805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864

Qy 241 GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260  
 |||||  
 Db 865 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGC 924

Qy 261 HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280  
 |||||  
 Db 925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC 984

Qy 281 LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300  
 |||||  
 Db 985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy 301 GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320  
 |||||  
 Db 1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA 1104

Qy 321 CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340  
 |||||  
 Db 1105 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCAGAGCAGCAGCTCC 1164

Qy 341 ProSer 342  
 |||||  
 Db 1165 CCCAGC 1170

# RESULT 3

AK056285

LOCUS AK056285 2703 bp mRNA linear PRI 20-JAN-2006

DEFINITION Homo sapiens cDNA FLJ31723 fis, clone NT2RI2006682, weakly similar to Human WW domain binding protein-1 mRNA.

ACCESSION AK056285

VERSION AK056285.1 GI:16551642

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS

Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,

Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,  
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,  
Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,  
Hara,H., Tanase,T.O., Nomura,Y., Togiya,S., Komai,F., Hara,R.,  
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,  
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,  
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,  
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,  
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,  
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,  
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,  
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,  
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,  
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,  
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,  
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,  
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,  
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,  
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,  
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

TITLE Complete sequencing and characterization of 21,243 full-length  
human cDNAs  
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)  
PUBMED 14702039  
REFERENCE 2  
AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,  
Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,  
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,  
Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,  
Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,  
Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,  
Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,  
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 2703)  
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.  
TITLE Direct Submission  
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:flj-cdna@nifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.

FEATURES  
source Location/Qualifiers  
1. .2703  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RI2006682"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RI2"  
/note="cloning vector: pME18SFL3  
mRNA from NT2 neuronal precursor cells treated 2-weeks  
mitotic inhibitor after 5-weeks retinoic acid (RA)  
induction.  
majorly NT2 neuron"

ORIGIN

Alignment Scores:  
Pred. No.: 9.67e-128 Length: 2703  
Score: 1901.00 Matches: 340  
Percent Similarity: 99.7% Conservative: 1  
Best Local Similarity: 99.4% Mismatches: 1  
Query Match: 99.6% Indels: 0  
DB: 5 Gaps: 0

US-10-729-895A-2 (1-342) x AK056285 (1-2703)

Qy	1	MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	145	ATGCCTTTCTCTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	204
Qy	21	SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	205	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	264
Qy	41	GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60
Db	265	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	324
Qy	61	CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn	80
Db	325	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	384
Qy	81	LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu	100
Db	385	CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTGTG	444
Qy	101	ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro	120
Db	445	CCAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA	504
Qy	121	ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla	140
Db	505	CCATACAGTGCCTTCCAGTACAGCAGCAGCAGTGTGCTCCACAGTGTGGCCCTGCA	564
Qy	141	GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro	160
Db	565	GGTGGCAGTCCCCGGGCATCGATCCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC	624
Qy	161	LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer	180
Db	625	TTGTCTGAGCCAGCAGAAGCAGCACAAAGCCCCAAGCATCGCTGACCCTGATCCCTCT	684
Qy	181	AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal	200
Db	685	GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	744
Qy	201	AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu	220
Db	745	GCTGGCTTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAGATGCAGAATGTAGGGAG	804
Qy	221	GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro	240
Db	805	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGAGCCT	864
Qy	241	GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly	260
Db	865	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGCAACCGGGGC	924
Qy	261	HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro	280
Db	925	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	984
Qy	281	LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys	300
Db	985	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGTGAGGAGGAAGGCCTCTGT	1044
Qy	301	GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla	320
Db	1045	CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA	1104
Qy	321	CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer	340
Db	1105	TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCAGAGCAGCAGCTCC	1164
Qy	341	ProSer	342
Db	1165	CCCAGC	1170

RESULT 4  
 CQ720290  
 LOCUS CQ720290 1002 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 6224 from Patent WO02068579.

ACCESSION CQ720290  
 VERSION CQ720290.1 GI:42281147  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 TITLE Kits, such as nucleic acid arrays, comprising a majority of  
 human exons or transcripts, for detecting expression and other uses  
 thereof  
 JOURNAL Patent: WO 02068579-A 6224 06-SEP-2002;  
 PE Corporation (NY) (US)  
 FEATURES Location/Qualifiers  
 source 1. 1002  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN

Alignment Scores:

Pred. No.:	1.77e-125	Length:	1002
Score:	1863.00	Matches:	333
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	97.6%	Indels:	0
DB:	2	Gaps:	0

US-10-729-895A-2 (1-342) x CQ720290 (1-1002)

```

Qy      10 AspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHisCys 29
      |||
Db      1 GATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACACTGC 60

Qy      30 CysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuValTrp 49
      |||
Db      61 TGTGGACAGTCTCAGTGCTGCACTACTACTATGAACTCTGGTGGTCTGGCTGGTGTGG 120

Qy      50 ThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHisArg 69
      |||
Db      121 ACCATCATCATCATCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCACC GC 180

Qy      70 LeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHisAsn 89
      |||
Db      181 CTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCCAACAT 240

Qy      90 TyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyrGlu 109
      |||
Db      241 TACTCAGCGCTGCCATTTTATTTTCTGAGTTTGGCCAACTATTTACTACCTCCTTATGAG 300

Qy      110 GluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGlnGln 129
      |||
Db      301 GAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACAGCAG 360

Qy      130 GlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAspPro 149
      |||
Db      361 CAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGATCCC 420

Qy      150 ThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSerThr 169
      |||
Db      421 ACCAGGGGATCCCAGGGGGACAGAGCAGCCCCTGTCTGAGCCAGCAGAGCAGCACA 480

Qy      170 ArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAlaThr 189
      |||
Db      481 AGACCCCAAGCATCGCTGACCTGATCCTCTGACCTACCACTTGACCGAGCAGCCACC 540

Qy      190 LysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspProGly 209
      |||
Db      541 AAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCGGGG 600

Qy      210 AlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSerGlu 229
      |||
Db      601 GCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCTGAA 660
  
```

Qy 230 HisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGlyAsp 249  
 Db 661 CACGGCGCACCCGACAGCAAAGAGACGCCTGGGAGACATCGCCGCTTACAGGTGAC 720

Qy 250 SerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGluPhe 269  
 Db 721 TCGGGCATTGAAGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTC 780

Qy 270 AsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHisVal 289  
 Db 781 AACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCATGTG 840

Qy 290 ArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArgGlu 309  
 Db 841 CGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGAGAG 900

Qy 310 ProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsnGlu 329  
 Db 901 CCTGGGCACCCGACCTGCCACGGCCGCCGCATGCCTGCTGCTGAACACCATCAACGAG 960

Qy 330 GlnAspSerProAsnSerGlnSerSerSerSerProSer 342  
 Db 961 CAGGACTCTCCCACTCCAGAGCAGCAGCTCCCCAGC 999

RESULT 5  
 CQ413042

LOCUS CQ413042 3095 bp DNA linear PAT 23-JAN-2004

DEFINITION Sequence 20113 from Patent WO0170979.

ACCESSION CQ413042

VERSION CQ413042.1 GI:41320823

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE

1 Lee, J. and Lillie, J.

TITLE Genes, compositions, kits, and method for identification,  
 assessment, prevention, and therapy of ovarian cancer

JOURNAL Patent: WO 0170979-A 20113 27-SEP-2001;

Millennium Pharmaceuticals, Inc. (US)

FEATURES

Location/Qualifiers

source 1. .3095

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	6.34e-125	Length:	3095
Score:	1863.00	Matches:	333
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	97.6%	Indels:	0
DB:	2	Gaps:	0

US-10-729-895A-2 (1-342) x CQ413042 (1-3095)

Qy 9 GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28  
 Db 109 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 168

Qy 29 CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48  
 Db 169 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 228

Qy 49 TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68  
 Db 229 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 288

Qy 69 ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis 88  
 Db 289 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 348

Qy 89 AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr 108

```

Db      349  |||||AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACTATTACTACCTCCTTAT 408
Qy      109  GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln 128
Db      409  |||||GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATAACAGTGCCTTCCAGCTACAG 468
Qy      129  GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp 148
Db      469  |||||CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 528
Qy      149  ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer 168
Db      529  |||||CCCACCAGGGGATCCCAGGGGGCACAGACAGCCCTTGTCTGAGCCAGCAGAAGCAGC 588
Qy      169  ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla 188
Db      589  |||||ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 648
Qy      189  ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro 208
Db      649  |||||ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 708
Qy      209  GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer 228
Db      709  |||||GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 768
Qy      229  GluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGly 248
Db      769  |||||GAACACGGCGACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTACAGGT 828
Qy      249  AspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGlu 268
Db      829  |||||GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 888
Qy      269  PheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHis 288
Db      889  |||||TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 948
Qy      289  ValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArg 308
Db      949  |||||GTGCGGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCAGCCCTCTGAGGAGCAGGCTCGA 1008
Qy      309  GluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsn 328
Db      1009 |||||GAGCCTGGGCACCCGCACCTGCCACGGCCGCCGCATGCCTGCTGCTGAACACCATCAAC 1068
Qy      329  GluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342
Db      1069 |||||GAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCCCCCAGC 1110

```

# RESULT 6

BD275013

LOCUS BD275013 4054 bp DNA linear PAT 17-JUL-2003

DEFINITION 50 Human Secreted Proteins.

ACCESSION BD275013

VERSION BD275013.1 GI:33084781

KEYWORDS JP 2002542766-A/13.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 4054)

AUTHORS Komatsoulis,G., Rosen,C.A. and Ruben,S.M.

TITLE 50 Human Secreted Proteins

JOURNAL Patent: JP 2002542766-A 13 17-DEC-2002;

Rosen et al

COMMENT OS Homo sapiens

PN JP 2002542766-A/13

PD 17-DEC-2002

PF 09-MAR-2000 JP 2000605624

PR 12-MAR-1999 US 60/124093,23-NOV-1999 US 60/166989 PI

george komatsoulis,craig a rosen,steven m ruben CC

FH Key Location/Qualifiers.

FEATURES Location/Qualifiers

source 1. .4054  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	1.46e-123	Length:	4054
Score:	1846.00	Matches:	333
Percent Similarity:	99.4%	Conservative:	0
Best Local Similarity:	99.4%	Mismatches:	1
Query Match:	96.7%	Indels:	1
DB:	2	Gaps:	0

US-10-729-895A-2 (1-342) x BD275013 (1-4054)

```
Qy      9  GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28
      |||
Db     107 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 166

Qy     29  CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48
      |||
Db     167 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 226

Qy     49  TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68
      |||
Db     227 TGGACCATCATCATCATCCTGAGTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 286

Qy     69  ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis 88
      |||
Db     287 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 346

Qy     89  AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProPro-Ty 108
      |||
Db     347 AATTACTCAGGCTGCCATTTTATTTAGGTTTTTGCCAACTATTTACTACCTCCTTTA 406

Qy    108  rGluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGl 128
      |||
Db     407 TGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATACAGTGCCTTCCAGCTACA 466

Qy    128  nGlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAs 148
      |||
Db     467 GCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGA 526

Qy    148  pProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSe 168
      |||
Db     527 TCCCACCAGGGGATCCCAGGGGGCAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAG 586

Qy    168  rThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAl 188
      |||
Db     587 CACAAGACCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGC 646

Qy    188  aThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPr 208
      |||
Db     647 CACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCC 706

Qy    208  oGlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSe 228
      |||
Db     707 GGGGGCCTTCTCGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTC 766

Qy    228  rGluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGl 248
      |||
Db     767 TGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGG 826

Qy    248  yAspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGl 268
      |||
Db     827 TGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGA 886

Qy    268  uPheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHi 288
      |||
Db     887 GTTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGTGCCA 946

Qy    288  sValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaAr 308
      |||
Db     947 TGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTAGTCCTCTGAGGAGCAGGCTCG 1006
```

Qy 308 gGluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAs 328  
 Db 1007 AGAGCCTGGGACCCGGACCTGCCACGGCCGCCGCATGCCTGCTGCTGAACACCATCAA 1066

Qy 328 nGluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342  
 Db 1067 CGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCCAGC 1109

# RESULT 7

BC026369

LOCUS BC026369 2241 bp mRNA linear ROD 28-JUL-2005

DEFINITION Mus musculus DNA segment, Chr 19, Wayne State University 162, expressed, mRNA (cDNA clone MGC:31104 IMAGE:4160029), complete cds.

ACCESSION BC026369

VERSION BC026369.1 GI:20071329

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2241)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

CONSRM Mammalian Gene Collection Program Team

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2241)

AUTHORS .

CONSRM NIH MGC Project

TITLE Direct Submission

JOURNAL Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 43 Row: c Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22122608.

FEATURES

source

Location/Qualifiers

1. .2241

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:31104 IMAGE:4160029"



```

/tissue_type="Liver, normal. 5 month old male mouse."
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
gene      1. .2241
          /gene="D19Wsu162e"
          /note="synonym: MGC31104"
          /db_xref="GeneID:226178"
          /db_xref="MGI:107577"
CDS       112. .1194
          /gene="D19Wsu162e"
          /codon_start=1
          /product="hypothetical protein LOC226178"
          /protein_id="AAH26369.1"
          /db_xref="GI:20071330"
          /db_xref="GeneID:226178"
          /db_xref="MGI:107577"
          /translation="MALLLLQALPSPLSVRAEPPQDKEACVGTNNQSYICDTGHCCGQ
SQCCNYYYELWFWLVWTVVILSCCCVCHHRAKHRLQAQQRQHEINLIAYREAHNY
SALPFYFREL PNSLLPPYEEV VNRPT PPPYS AFQLQQQQQLL PPPQGGPPGGSP
GTD PPPQGSQGAQSSPLSGPSRSSTRPPSVADPQSP EVPTDREATKASGTESGSPMAG
HGELDPGAFLDQDSECKEELLKDSR SERGGVSPDSEDKT PGRHRRFTGDSGIEVCVN
RGHHDDDLKEFNTLIDDALDGLDFDCDSCHVRPPVDEEEGLCLSEGGQAREHGHPHLP
RPPACLLLNTINEQDSPNSQHS GSPS"

```

# ORIGIN

## Alignment Scores:

Pred. No.:	2.36e-107	Length:	2241
Score:	1618.00	Matches:	296
Percent Similarity:	88.4%	Conservative:	15
Best Local Similarity:	84.1%	Mismatches:	31
Query Match:	84.8%	Indels:	10
DB:	6	Gaps:	4

US-10-729-895A-2 (1-342) x BC026369 (1-2241)

```

QY      1 MetProPheLeuLeuGlyLeuArg-----GlnAspLysGluAlaCysValGly 16
      :::|||   |||   :::|||   |||||
Db      136 CTGCCAGCCCTGTCTAGTGAGGGCTGAGCCCCCGCAGGATAAGGAAGCCTGTGTGGGT 195

QY      17 ThrAsnAsnGlnSerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCys 36
      |||||
Db      196 ACCAACAATCAAGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGC 255

QY      37 AsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSer 56
      |||||
Db      256 AACTACTACTATGAACCTCTGGTGGTTCTGGCTGGTGTGGACCGTCGTCATCATTCTGAGC 315

QY      57 CysCysCysValCysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGln 76
      |||||
Db      316 TGCTGCTGTGTCTGCCACCACCGCCGAGCCAAGCACC GCCTTCAGGCTCAGCAGCGGCAA 375

QY      77 HisGluIleAsnLeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyr 96
      |||||
Db      376 CATGAAATCAACCTGATCGCTTACCGGGAAGCCCACTACTCAGCGTTGCCGTTTAC 435

QY      97 PheArgPheLeuProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProPro 116
      |||||
Db      436 TTCAGGTTTTTGCCAAACTCTTGTCTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCA 495

QY      117 ThrProProProProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeu----- 133
      |||||
Db      496 ACTCCTCCCCACCGTACAGTGCCTTCCAGCTCCAACAGCAGCAGCAGCTGCTGCCACCT 555

QY      134 ProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp---ProThrArgGly 152
      |||||
Db      556 CCTCCTCAGGTTGGCCCTCCAGGTGGCAGCCCCCAGGCACTGACCCGCCGCCAGGGC 615

QY      153 SerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSerThrArgProPro 172
      |||||
Db      616 TCCAGGGAGCGCAGAGCAGTCTTTGTCCGACCGAGCCGAAGCAGCACAAAGACCTCCA 675

QY      173 SerIleAlaAspProAspProSerAspLeuProValAspArgAlaAlaThrLysAlaPro 192
      |||||
Db      676 AGTGTGCGAGATCCTCAGTCCCCTGAAGTGCCCACTGACCGGGAAGCCACCAAGCCTCT 735

```





12: gb\_gss2:\*  
13: gb\_gss3:\*  
14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query				DB ID	Description
	Score	Match	Length	DB		
1	1847.5	96.8	2792	6	CR859701	CR859701 Pongo pyg
2	1651	86.5	2603	6	AK052400	AK052400 Mus muscu
3	1651	86.5	2702	6	AK078491	AK078491 Mus muscu
4	1651	86.5	3871	6	AK142695	AK142695 Mus muscu
5	1616	84.7	3893	6	AK154686	AK154686 Mus muscu
6	1593	83.4	3873	6	AK087487	AK087487 Mus muscu
7	1576.5	82.6	3955	6	AK169294	AK169294 Mus muscu
8	1304	68.3	735	9	DB293556	DB293556 DB293556
9	1271	66.6	800	8	CX163170	CX163170 HESC2_2_F
10	1093.5	57.3	852	9	DN935380	DN935380 AGENCOURT
11	1093	57.3	876	9	DN948780	DN948780 AGENCOURT
12	1086	56.9	712	8	CN427264	CN427264 170005321
13	1081	56.6	796	5	CF737347	CF737347 UI-M-HD0-
14	1069	56.0	680	8	CV868359	CV868359 PDUTs1054
15	1067	55.9	587	9	DA190287	DA190287 DA190287
16	1059	55.5	762	4	CA316710	CA316710 UI-M-FW0-
17	1046	54.8	575	9	DA553446	DA553446 DA553446
18	1022	53.5	581	3	BP242855	BP242855 BP242855
19	1015.5	53.2	870	9	DN078935	DN078935 JGI_CABD1
20	1010	52.9	550	9	DB283277	DB283277 DB283277
21	994	52.1	550	9	DA576326	DA576326 DA576326
22	993	52.0	565	3	BP221694	BP221694 BP221694
23	982.5	51.5	844	8	CV118164	CV118164 AGENCOURT
24	982	51.4	566	9	DA806792	DA806792 DA806792
25	976	51.1	541	9	DB133804	DB133804 DB133804
26	974.5	51.0	888	9	DN076706	DN076706 JGI_CABD9
27	973.5	51.0	882	9	CX964026	CX964026 JGI_CAAO1
28	969	50.8	921	9	DN076178	DN076178 JGI_CABD9
29	951	49.8	625	4	CA334353	CA334353 NISC_1s08
30	946	49.6	560	9	DA756667	DA756667 DA756667
31	944	49.4	553	9	DA760493	DA760493 DA760493
32	905	47.4	488	9	DA596718	DA596718 DA596718
33	901	47.2	563	9	DA321605	DA321605 DA321605
34	899	47.1	716	3	BU701616	BU701616 UI-M-FIO-
35	882	46.2	818	3	BQ769540	BQ769540 UI-M-FIO-
36	874	45.8	525	9	CX597272	CX597272 CT020017B
37	869.5	45.5	1006	3	BQ054184	BQ054184 AGENCOURT
38	868	45.5	624	3	BU053344	BU053344 UI-M-FCO-
39	861	45.1	537	9	DB129660	DB129660 DB129660
40	858	44.9	652	7	BB658981	BB658981 BB658981
41	847	44.4	559	1	AI786533	AI786533 uj17b01.y
42	844	44.2	580	9	DA823881	DA823881 DA823881
43	831	43.5	820	8	CN094623	CN094623 EC2BAA1AC
44	830	43.5	703	4	BY748360	BY748360 BY748360
45	820	43.0	565	9	DA196239	DA196239 DA196239

#### ALIGNMENTS

RESULT 1  
CR859701  
LOCUS CR859701 2792 bp mRNA linear HTC 12-NOV-2004  
DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp469N2325 (from clone DKFZp469N2325).  
ACCESSION CR859701  
VERSION CR859701.1 GI:55730274  
KEYWORDS HTC.  
SOURCE Pongo pygmaeus (orangutan)  
ORGANISM Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Pongo.  
REFERENCE 1 (bases 1 to 2792)  
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,

Fobo,G., Han,M. and Wiemann,S.  
 CONSRTM The German cDNA Consortium  
 TITLE Direct Submission  
 JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.  
 This clone (DKFZp469N2325) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:  
 http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469N2325  
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.  
 FEATURES Location/Qualifiers  
 source 1. .2792  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp469N2325"  
 /tissue\_type="kidney"  
 /clone\_lib="469 (synonym: pkid1). Vector pSport1\_Sfi; host DH10B; sites SfilA + SfilB"  
 /dev\_stage="adult"  
 /note="hypothetical protein (Mus musculus), differentially spliced"  
 gene 1. .2792  
 /gene="DKFZp469N2325"  
 CDS 43. .360  
 /gene="DKFZp469N2325"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAH91860.1"  
 /db\_xref="GI:55730275"  
 /translation="MERRRLLGGMALLLLQALPSPLSARAEPQLFGGASCGSLSCVW LTTFALLVQYWGFQGRIRKPVWVPTIKATSVTQDTAVDSLAAATTTMNSGGSGWCGPS SSS"

#### ORIGIN

#### Alignment Scores:

Pred. No.:	6.11e-133	Length:	2792
Score:	1847.50	Matches:	331
Percent Similarity:	98.8%	Conservative:	5
Best Local Similarity:	97.4%	Mismatches:	3
Query Match:	96.8%	Indels:	1
DB:	6	Gaps:	1

US-10-729-895A-2 (1-342) x CR859701 (1-2792)

```

Qy      4  LeuLeuGlyLeuArg---GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyr 22
      :::::::::::  |||
Db      201 GTACTGGGGGTTTCAGGGCAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTAC 260

Qy      23  IleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeu 42
      |||
Db      261 ATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAATC 320

Qy      43  TrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysValCysHis 62
      |||
Db      321 TGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTTGCCAC 380

Qy      63  HisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIle 82
      |||
Db      381 CACCGCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATC 440

Qy      83  AlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsn 102
      |||
Db      441 GCCTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTATTTAGGTTTTCGCAAAC 500

Qy      103 TyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProProTyr 122
      |||
Db      501 TATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATAC 560

Qy      123 SerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGly 142
      |||

```

Db 561 AGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGC 620

Qy 143 SerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSer 162  
 |||||:::|||||

Db 621 AGTCCCCCGGCGTTGATCCACCAAGGGGATCCAGGGGGCAGAGCAGCCCTTGTCT 680

Qy 163 GluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSerAspLeu 182  
 |||||:::|||||

Db 681 GGGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCTGAGCCCTCTGACCTA 740

Qy 183 ProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerValAlaGly 202  
 |||||:::|||||

Db 741 CCAGTTGAACGAGCAGCCACCAAGCCCCGGGGATGGAGCCAGTGGCTCTGTGGCTGGC 800

Qy 203 LeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeu 222  
 |||||

Db 801 CTGGGGGAGCTGGACCCGGGGGCTTCTGGACAAGGATGCAGAATGTAGGGAGGAGCTG 860

Qy 223 LeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrProGlyArg 242  
 |||||

Db 861 CTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGA 920

Qy 243 HisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGlyHisHis 262  
 |||||

Db 921 CATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGCAACCGGGGCCACCAT 980

Qy 263 AspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAsp 282  
 |||||

Db 981 GACGATGACCTCAAAGAGTTCAACACGCTCATCGATGATGCTCTGGATGGGCCCTGGAC 1040

Qy 283 PheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCysGlnSer 302  
 |||||

Db 1041 TTCTGCGATAGCTGCCATGTGCGCCCCCTGGTGACGAGGAGGAAGGCCTCTGCCAGCCC 1100

Qy 303 SerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAlaCysLeu 322  
 |||||

Db 1101 TCCGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCGCATGCCTG 1160

Qy 323 LeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342  
 |||||

Db 1161 CTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCAGAGCAGCAGCTCCCCCAGC 1220

RESULT 2

AK052400

LOCUS AK052400 2603 bp mRNA linear HTC 02-SEP-2005

DEFINITION Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched library, clone:D430003C10 product:hypothetical Proline-rich region/Cytochrome c family heme-binding site containing protein, full insert sequence.

ACCESSION AK052400

VERSION AK052400.1 GI:26342634

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1

AUTHORS Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 PUBMED 10349636

REFERENCE

2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 PUBMED 11042159

REFERENCE

3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.; Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium, the RIKEN Genome Exploration Research Group  
Phase I and II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6

AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group  
(Genome Network Core Team) and the FANTOM Consortium.

TITLE Antisense Transcription in the Mammalian Transcriptome

JOURNAL Science 309, 1564-1566 (2005)

REFERENCE 7

AUTHORS The FANTOM Consortium, Riken Genome Exploration Research Group and  
Genome Science Group (Genome Network Project Core Group).

TITLE The Transcriptional Landscape of the Mammalian Genome

JOURNAL Science 309, 1559-1563 (2005)

REFERENCE 8 (bases 1 to 2603)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp;  
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.jp/  
URL:http://fantom.gsc.riken.jp/.

FEATURES

source 1. .2603  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:D430003C10"  
/db\_xref="taxon:10090"  
/clone="D430003C10"  
/tissue\_type="lung"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="13 days embryo"

CDS 143. .1189  
/note="unnamed protein product; hypothetical Proline-rich  
region/Cytochrome c family heme-binding site containing  
protein (InterPro|IPR000694 InterPro|IPR000345, evidence:  
InterPro)  
putative"  
/codon\_start=1  
/protein\_id="BAC34974.1"  
/db\_xref="GI:26342635"  
/translation="MPFLWGLRQDKEACVGTNNQSYICDTGHCCGQSQCNNYYELWW"

```

FWLVWTVVVIILSCCVCHHRAKHRLQAQQROHEINLIAYREAHNYSALPFYFRFLPN
SLPPYEEVVRNPPTPPPPYSAFQLQQQQQLPPPPQGGPPGSPPGADPPPPQGSQGA
QSSPLSGPSRSSTRPPSVADPQSPVPTDREATKASGTESGSPMAGHGELDPGAFLDQ
DSECKEELLKDSRSEKGVSPDSEKTPGRHRRFTGDSGIEVCVNRGHDDDLKEFN
TLIDDALDGLDFCDSCHVRPPVDEEEGLCLSSEGQAREHGPHLPRPPACLLNTIN
EQDSPNSQHSQSPS"
polyA_signal      2584..2589
                  /note="putative"
polyA_site        2603
                  /note="putative"
ORIGIN
Alignment Scores:
Pred. No.:        1.13e-117      Length:      2603
Score:            1651.00        Matches:      300
Percent Similarity: 89.9%        Conservative: 13
Best Local Similarity: 86.2%      Mismatches:   29
Query Match:      86.5%          Indels:        6
DB:               6              Gaps:          3

US-10-729-895A-2 (1-342) x AK052400 (1-2603)

Qy      1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
      |||
Db      143 ATGCCTTTCTCTGCGGGCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 202

Qy      21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40
      |||
Db      203 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 262

Qy      41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60
      |||
Db      263 GAACTCTGGTGGTTCTGGCTGGTGTGGACCGTCGTCATCATTCTGAGCTGCTGCTGTGTC 322

Qy      61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80
      |||
Db      323 TGCCACCACCGCCGAGCCAGCCAGCCGCTTCAGGCTCAGCAGCGGCAACATGAAATCAAC 382

Qy      81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100
      |||
Db      383 CTGATCGCTTACCGGAAGCCACAACTACTCAGCGTTGCCGTTTACTTCAAGTTTGTG 442

Qy      101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120
      |||
Db      443 CCAAACTCTTTGCTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA 502

Qy      121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeu-----ProProGlnCys 137
      |||
Db      503 CCGTACAGTGCCTTCCAGCTCCAACAGCAGCAGCAGCTGCTGCCACCTCCTCCTCAGGGT 562

Qy      138 GlyProAlaGlyGlySerProProGlyIleAsp---ProThrArgGlySerGlnGlyAla 156
      |||
Db      563 GGCCCTCCAGGTGGCAGCCCCCAGGCGCTGACCCGCGCCCCAGGGCTCCAGGGAGCG 622

Qy      157 GlnSerSerProLeuSerGluProSerArgSerSerThrArgProProSerIleAlaAsp 176
      |||
Db      623 CAGAGCAGTCTTTGTCCGACCGAGCCGAAGCAGCACAAGACCTCCAAGTGTGCAGAT 682

Qy      177 ProAspProSerAspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluPro 196
      |||
Db      683 CCTCAGTCCCTGAAGTGCCCACTGACCGGAAGCCACCAAAGCCTCTGGGACGGAGTCT 742

Qy      197 SerGlySerValAlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAla 216
      |||
Db      743 GGTAGCCCATGGCAGGCCACGGGAGCTGGACCGGGCGCCTTCTGGACAGGATTCC 802

Qy      217 GluCysArgGluGluLeuLeuLysAspAspSerSerGluHisGly-----AlaProAsp 234
      |||
Db      803 GAGTGCAAGGAGGAGCTACTGAAAGATTGCGGTTCTGAGCGTGGCGCGTGTCCCTGAC 862

Qy      235 SerLysGluLysThrProGlyArgHisArgArgPheThrGlyAspSerGlyIleGluVal 254
      |||
Db      863 AGTGAAGACAAGACCCCGGAGGATCGCCGCTTCACAGGTGACTCAGGCATTGAGGTG 922

Qy      255 CysValCysAsnArgGlyHisHisAspAspAspLeuLysGluPheAsnThrLeuIleAsp 274
      |||

```



Db 923 TGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTTAACACGCTCATAGAT 982

QY 275 AspAlaLeuAspGlyProLeuAspPheCysAspSerCysHisValArgProProGlyAsp 294  
 |||

Db 983 GATGCTCTGGACGGGCCCTGGACTTCTGTGACAGCTGTCATGTACGGCCTCCTGTTGAT 1042

QY 295 GluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArgGluProGlyHisProHis 314  
 |||

Db 1043 GAAGAGGAAGGTCTCTGCCTGTCTCGGAGGGTCAGGCTCGGGAGCATGGGCACCCCCAC 1102

QY 315 LeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsn 334  
 |||

Db 1103 CTGCCACGGCCGCTGCGTGTCTGTCTGCTAAACACCATCAATGAGCAGGACTCTCCAAAT 1162

QY 335 SerGlnSerSerSerSerProSer 342  
 |||

Db 1163 TCCCAGCACAGTGGCTCCCCGAGC 1186

# RESULT 3

AK078491

LOCUS AK078491 2702 bp mRNA linear HTC 02-SEP-2005

DEFINITION Mus musculus 12 days embryo female mullerian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6820415N12 product:hypothetical Proline-rich region/Cytochrome c family heme-binding site containing protein, full insert sequence.

ACCESSION AK078491

VERSION AK078491.1 GI:26347312

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

## REFERENCE

1  
 AUTHORS Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 PUBMED 10349636

## REFERENCE

2  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 PUBMED 11042159

## REFERENCE

3  
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 PUBMED 11076861

## REFERENCE

4  
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)

## REFERENCE

5  
 AUTHORS The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)

## REFERENCE

6  
 AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.  
 TITLE Antisense Transcription in the Mammalian Transcriptome  
 JOURNAL Science 309, 1564-1566 (2005)

## REFERENCE

7  
 AUTHORS The FANTOM Consortium, Riken Genome Exploration Research Group and

Genome Science Group (Genome Network Project Core Group).  
**TITLE** The Transcriptional Landscape of the Mammalian Genome  
**JOURNAL** Science 309, 1559-1563 (2005)  
**REFERENCE** 8 (bases 1 to 2702)  
**AUTHORS** Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
 Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
 Muramatsu,M. and Hayashizaki,Y.

**TITLE** Direct Submission  
**JOURNAL** Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,  
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)

**COMMENT** cdna library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/.

**FEATURES**  
 source Location/Qualifiers  
 1. .2702  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:6820415N12"  
 /db\_xref="taxon:10090"  
 /clone="6820415N12"  
 /sex="female"  
 /tissue\_type="mullerian duct includes surrounding region"  
 /clone\_lib="RIKEN full-length enriched mouse cdna library"  
 /dev\_stage="12 days embryo"  
 CDS 154. .1200  
 /note="unnamed protein product; hypothetical Proline-rich  
 region/Cytochrome c family heme-binding site containing  
 protein (InterPro|IPR000694 InterPro|IPR000345, evidence:  
 InterPro)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC37305.1"  
 /db\_xref="GI:26347313"  
 /translation="MPFLWGLRQDKEACVGTNNQSYICDTGHCCGQSQCENYYELWW  
 FWLWVTVVILSCCVCHHRAKHLQAQQRQHEINLIAYREAHNYSALPFYFRFLPN  
 SLLPPYEEVNRPTPPPPYSAFQLQQQQQLLPPPPQGGPPGGSPGADPPPGSQGA  
 QSSPLSGPSRSSTRPPSVADPQSPEVPTDREATKASGTESGSPMAGHGELDPGAFLDQ  
 DSECKEELKDSRERGGVSPDSEDKTPGRHRRFTGDSGIEVCVNCNRGHDDDLKEFN  
 TLIDDALDGLDFCDSCHVRPVPDEEEGLCLSSSEGQAREHGPHLPRPPACLLNTIN  
 EQDSPNSQHSQSPS"

#### ORIGIN

#### Alignment Scores:

Pred. No.:	1.19e-117	Length:	2702
Score:	1651.00	Matches:	300
Percent Similarity:	89.9%	Conservative:	13
Best Local Similarity:	86.2%	Mismatches:	29
Query Match:	86.5%	Indels:	6
DB:	6	Gaps:	3

US-10-729-895A-2 (1-342) x AK078491 (1-2702)

Qy 1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20  
 |||||  
 Db 154 ATGCCCTTCCTCTGGGGGCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 213

Qy 21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40  
 Db 214 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 273

Qy 41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60  
 Db 274 GAACTCTGGTGGTTCTGGCTGGTGGACCGTCGTCATCATTCTGAGCTGCTGCTGTGTC 333

Qy 61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80  
 Db 334 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCTCAGCAGCGGCAACATGAAATCAAC 393

Qy 81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100  
 Db 394 CTGATCCTTACCAGGGAAGCCACAACTACTCAGCGTTGCCGTTTACTTCAGGTTTTG 453

Qy 101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120  
 Db 454 CCAAATCTTTGCTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAATCCTCCCCA 513

Qy 121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeu-----ProProGlnCys 137  
 Db 514 CCGTACAGTGCCTTCAGCTCCAACAGCAGCAGCAGCTGCTGCCACCTCCTCCTCAGGGT 573

Qy 138 GlyProAlaGlyGlySerProProGlyIleAsp---ProThrArgGlySerGlnGlyAla 156  
 Db 574 GGCCCTCCAGGTGGCAGCCCCCAGGCGCTGACCCGCCGCCAGGGCTCCAGGGAGCG 633

Qy 157 GlnSerSerProLeuSerGluProSerArgSerSerThrArgProProSerIleAlaAsp 176  
 Db 634 CAGAGCAGTCCTTTGTCCGACCGAGCCGAAGCAGCACAAGACCTCCAAGTGTGCGAGAT 693

Qy 177 ProAspProSerAspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluPro 196  
 Db 694 CCTCAGTCCCTGAAGTGCCCACTGACCGGGAAGCCACCAAAGCCTCTGGGACGGAGTCT 753

Qy 197 SerGlySerValAlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAla 216  
 Db 754 GGTAGCCCCATGGCCGGCCACGGGAGCTGGACCCGGGCGCCTTCTGGACAGGATTCC 813

Qy 217 GluCysArgGluGluLeuLeuLysAspAspSerSerGluHisGly-----AlaProAsp 234  
 Db 814 GAGTGCAAGGAGGAGCTACTGAAAGATTGCGCTTCTGAGCGTGGCGCGGTGCCCTGAC 873

Qy 235 SerLysGluLysThrProGlyArgHisArgArgPheThrGlyAspSerGlyIleGluVal 254  
 Db 874 AGTGAAGACAAGACCCCGGAGGATCGCCGCTTCACAGGTGACTCAGGCATTGAGGTG 933

Qy 255 CysValCysAsnArgGlyHisHisAspAspAspLeuLysGluPheAsnThrLeuIleAsp 274  
 Db 934 TGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTTAACACGCTCATAGAT 993

Qy 275 AspAlaLeuAspGlyProLeuAspPheCysAspSerCysHisValArgProProGlyAsp 294  
 Db 994 GATGCTCTGGACGGGCCCCTGGACTTCTGTGACAGCTGTCATGTACGGCCTCCTGTTGAT 1053

Qy 295 GluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArgGluProGlyHisProHis 314  
 Db 1054 GAAGAGGAAGGTCTCTGCCTGTCTCGGAGGGTCAGGCTCGGGAGCATGGGCACCCCCAC 1113

Qy 315 LeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsn 334  
 Db 1114 CTGCCACGGCCGCCTGCGTGTCTGCTGCTAAACACCATCAATGAGCAGGACTCTCCAAAT 1173

Qy 335 SerGlnSerSerSerSerProSer 342  
 Db 1174 TCCAGCACAGTGGCTCCCCGAGC 1197

RESULT 4

AK142695

LOCUS AK142695 3871 bp mRNA linear HTC 21-SEP-2005

DEFINITION Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830007022 product:hypothetical protein, full insert sequence.

ACCESSION

AK142695

VERSION

AK142695.1 GI:74216502

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

CONSRM RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

PUBMED 11217851

REFERENCE 5

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

CONSRTM FANTOM Consortium  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420 (6915), 563-573 (2002)  
PUBMED 12466851  
REFERENCE 6  
AUTHORS Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzius,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavolan,M., Davis,M.J., Wilming,L.G., Aidinis,V., Allen,J.E., Ambesi-Impiombato,A., Apweiler,R., Aturaliya,R.N., Bailey,T.L., Bansal,M., Baxter,L., Beisel,K.W., Bersano,T., Bono,H., Chalk,A.M., Chiu,K.P., Choudhary,V., Christoffels,A., Clutterbuck,D.R., Crowe,M.L., Dalla,E., Dalrymple,B.P., de Bono,B., Della Gatta,G., di Bernardo,D., Down,T., Engstrom,P., Fagiolini,M., Faulkner,G., Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E., Gustincich,S., Harbers,M., Hayashi,Y., Hensch,T.K., Hirokawa,N., Hill,D., Huminicki,L., Iacono,M., Ikeo,K., Iwama,A., Ishikawa,T., Jakt,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitamura,H., Kitano,H., Kollias,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K., Kurochkin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liuni,S., McWilliam,S., Madan Babu,M., Madera,M., Marchionni,L., Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Mottagui-Tabar,S., Mulder,N., Nakano,N., Nakauchi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J., Pavesi,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Rost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C., Schonbach,C., Sekiguchi,K., Semple,C.A., Seno,S., Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sinclair,B., Sperling,S., Stupka,E., Sugiura,K., Sultana,R., Takenaka,Y., Taki,K., Tammouja,K., Tan,S.L., Tang,S., Taylor,M.S., Tegner,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,E., Verardo,R., Wei,C.L., Yagi,K., Yamanishi,H., Zabarovsky,E., Zhu,S., Zimmer,A., Hide,W., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T., Brusic,V., Quackenbush,J., Wahlestedt,C., Mattick,J.S., Hume,D.A., Kai,C., Sasaki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M., Suzuki,M., Aoki,J., Arakawa,T., Iida,J., Imamura,K., Itoh,M., Kato,T., Kawaji,H., Kawagashira,N., Kawashima,T., Kojima,M., Kondo,S., Konno,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessy,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watahiki,A., Okamura-Oho,Y., Suzuki,H., Kawai,J. and Hayashizaki,Y.

CONSRTM FANTOM Consortium  
TITLE The transcriptional landscape of the mammalian genome  
JOURNAL Science 309 (5740), 1559-1563 (2005)  
PUBMED 16141072  
REFERENCE 7  
AUTHORS Katayama,S., Tomaru,Y., Kasukawa,T., Waki,K., Nakanishi,M., Nakamura,M., Nishida,H., Yap,C.C., Suzuki,M., Kawai,J., Suzuki,H., Carninci,P., Hayashizaki,Y., Wells,C., Frith,M., Ravasi,T., Pang,K.C., Hallinan,J., Mattick,J., Hume,D.A., Lipovich,L., Batalov,S., Engstrom,P.G., Mizuno,Y., Faghihi,M.A., Sandelin,A., Chalk,A.M., Mottagui-Tabar,S., Liang,Z., Lenhard,B. and Wahlestedt,C.

CONSRTM RIKEN Genome Exploration Research Group  
TITLE Antisense transcription in the mammalian transcriptome  
JOURNAL Science 309 (5740), 1564-1566 (2005)  
PUBMED 16141073  
REFERENCE 8 (bases 1 to 3871)  
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hori,F., Iida,J., Imamura,K., Imotani,K., Itoh,M., Kanagawa,S., Kawai,J., Kojima,M., Konno,H., Murata,M., Nakamura,M., Ninomiya,N., Nishiyori,H., Nomura,K., Ohno,M., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shiraki,T., Tagami,M., Tagami,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission  
JOURNAL Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,  
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/.

FEATURES

source Location/Qualifiers  
 1. .3871  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:D830007022"  
 /db\_xref="taxon:10090"  
 /clone="D830007022"  
 /tissue\_type="heart"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="16 days neonate"

CDS  
 154..1200  
 /note="unnamed protein product; hypothetical protein  
 (evidence: CDS)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAE25164.1"  
 /db\_xref="GI:74216503"  
 /translation="MPFLWGLRQDKEACVGTNNQSYICDTGHCCGQSQCNCYYELWW  
 FWLVWTVVILSCCCVCHRRRAKHLQAQQRQHEINLIAYREAHNYSALPFYFRFLPN  
 SLLPPYEEVVRNPPTPPPPYSAFQLQQQQLLPPPPQGGPPGGSPPGADPPPPQGSQGA  
 QSSPLSGPSRSSTRPPSVADPQSPEVPTDREATKASGTESGSPMAGHGELDPGAFLDQ  
 DSECKEELLKDSRSEGGVSPDSEDKTPGRHRRFTGDSGIEVCVCNRGHHDDDLKEFN  
 TLIDDALDGLDFCDSCSHVRPVDDEEGLCLSSSEGQAREHGHPLPRPPACLLNTIN  
 EQDSPNSQHSQSPS"

polyA\_signal 3846..3851  
 /note="putative"

polyA\_site 3871  
 /note="putative"

Alignment Scores:

Pred. No.:	1.89e-117	Length:	3871
Score:	1651.00	Matches:	300
Percent Similarity:	89.9%	Conservative:	13
Best Local Similarity:	86.2%	Mismatches:	29
Query Match:	86.5%	Indels:	6
DB:	6	Gaps:	3

US-10-729-895A-2 (1-342) x AK142695 (1-3871)

Qy	1	MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	154	ATGCCTTCTCTGGGGCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	213
Qy	21	SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	214	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	273
Qy	41	GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60
Db	274	GAAGCTCTGGTGGTTCTGGCTGGTGTGGACCGTCGTCATCATTCTGAGCTGCTGTGTGC	333
Qy	61	CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn	80
Db	334	TGCCACCACCGCCGAGCCAAGCACCCTTCAGGCTCAGCAGCGGCAACATGAAATCAAC	393
Qy	81	LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu	100
Db	394	CTGATCGCTTACCGGAAGCCCACTACTCAGCGTTGCCGTTTACTTCAGGTTTTTG	453
Qy	101	ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro	120
Db	454	CCAAACTCTTTGCTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA	513

Qy 121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeu-----ProProGlnCys 137  
 Db 514 CCGTACAGTGCCTTCCAGCTCCAACAGCAGCAGCAGCTGCTGCCACCTCCTCCTCAGGGT 573

Qy 138 GlyProAlaGlyGlySerProProGlyIleAsp---ProThrArgGlySerGlnGlyAla 156  
 Db 574 GGCCCTCCAGGTGGCAGCCCCCAGGCGCTGACCCGCCGCCCCAGGGCTCCCAGGGAGCG 633

Qy 157 GlnSerSerProLeuSerGluProSerArgSerSerThrArgProProSerIleAlaAsp 176  
 Db 634 CAGAGCAGTCTTGTCCGACCGAGCCGAAGCAGCACAAGACCTCCAAGTGTGCGAGAT 693

Qy 177 ProAspProSerAspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluPro 196  
 Db 694 CCTCAGTCCCCTGAAGTGCCACTGACCGGGAAGCCACCAAAGCCTCTGGGACGGAGTCT 753

Qy 197 SerGlySerValAlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAla 216  
 Db 754 GGTAGCCCATGGCCGCCACGGGAGCTGGACCGGGCGCCTTCTGGACCAGGATTCC 813

Qy 217 GluCysArgGluGluLeuLeuLysAspAspSerSerGluHisGly-----AlaProAsp 234  
 Db 814 GAGTGCAGGAGGAGCTACTGAAAGATTGCGTTCTGAGCGTGGCGCGTGTCCCTGAC 873

Qy 235 SerLysGluLysThrProGlyArgHisArgArgPheThrGlyAspSerGlyIleGluVal 254  
 Db 874 AGTGAAGACAAGACCCCGGAGGATCGCCGCTTACAGGTGACTCAGGCATTGAGGTG 933

Qy 255 CysValCysAsnArgGlyHisHisAspAspAspLeuLysGluPheAsnThrLeuIleAsp 274  
 Db 934 TGTGTGTGCAACCGGGCCACCATGACGATGACCTCAAAGAGTTTAACACGCTCATAGAT 993

Qy 275 AspAlaLeuAspGlyProLeuAspPheCysAspSerCysHisValArgProProGlyAsp 294  
 Db 994 GATGCTCTGGACGGGCCCCTGGACTTCTGTGACAGCTGTCATGTACGGCCTCCTGTTGAT 1053

Qy 295 GluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArgGluProGlyHisProHis 314  
 Db 1054 GAAGAGGAAGGTCTCTGCCTGTCTCGGAGGGTCAGGCTCGGGAGCATGGGCACCCCCAC 1113

Qy 315 LeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsn 334  
 Db 1114 CTGCCACGGCCGCTGCGTGTCTGCTGCTAAACACCATCAATGAGCAGGACTCTCCAAAT 1173

Qy 335 SerGlnSerSerSerSerProSer 342  
 Db 1174 TCCCAGCACAGTGGCTCCCCGAGC 1197

# RESULT 5

AK154686

LOCUS AK154686 3893 bp mRNA linear HTC 21-SEP-2005

DEFINITION Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F630103G18 product:hypothetical protein, full insert sequence.

ACCESSION AK154686

VERSION AK154686.1 GI:74185772

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

## REFERENCE

1  
 AUTHORS Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 PUBMED 10349636

2  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 PUBMED 11042159

3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,  
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
 Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 PUBMED 11076861  
 REFERENCE 4  
 AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
 Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
 Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,  
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
 Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
 Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,  
 Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,  
 Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,  
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
 Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
 Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,  
 Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,  
 Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
 Marchionni,L., Mashima,J., Mazzairelli,J., Mombaerts,P., Nordone,P.,  
 Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
 Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
 Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,  
 Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.  
 and Hayashizaki,Y.  
 CONSRTM RIKEN Genome Exploration Research Group Phase II Team and the

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006



# SCORE Search Results Details for Application 10729895 and Search Result 20070112\_180212\_us-10-729-895a-2.rni.

<a href="#">Score Home</a>	<a href="#">Retrieve Application</a>	<a href="#">SCORE System</a>	<a href="#">SCORE</a>	<a href="#">Comments /</a>
<a href="#">Page</a>	<a href="#">List</a>	<a href="#">Overview</a>	<a href="#">FAQ</a>	<a href="#">Suggestions</a>

This page gives you Search Results detail for the Application 10729895 and Search Result 20070112\_180212\_us-10-729-895a-2.rni.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2007 Biocceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2007, 05:14:59 ; Search time 245 Seconds  
(without alignments).  
3917.873 Million cell updates/sec

Title: US-10-729-895A-2  
Perfect score: 1909  
Sequence: 1 MPFLLGLRQDKEACVGTNNQ.....LLNTINEQDSPNSQSSSSPS 342

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abss/ABSSWEB\_spool/US10729895/runat\_12012007\_180212\_7416/app\_query.fasta\_1  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abss03h -USER=US10729895\_CGN\_1\_1\_204 @runat\_12012007\_180212\_7416  
-NCPU=6 -ICPU=3 -NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1\_COMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5\_COMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A\_COMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B\_COMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7\_COMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H\_COMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS\_COMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP\_COMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE\_COMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query				ID	Description
	Score	Match	Length	DB		
1	1901	99.6	2703	4	US-10-094-749-823	Sequence 823, App
2	469.5	24.6	2042	4	US-10-094-749-344	Sequence 344, App
3	462	24.2	399	3	US-09-513-999C-274	Sequence 274, App
4	462	24.2	399	3	US-09-471-276-82	Sequence 82, Appl
5	452	23.7	1292	3	US-09-247-155-182	Sequence 182, App
6	452	23.7	1292	3	US-09-903-190-182	Sequence 182, App
7	329.5	17.3	861	3	US-09-247-155-83	Sequence 83, Appl
8	329.5	17.3	861	3	US-09-903-190-83	Sequence 83, Appl
9	253.5	13.3	489	3	US-09-513-999C-3746	Sequence 3746, Ap
10	150.5	7.9	2196	3	US-09-949-016-5786	Sequence 5786, Ap
11	145.5	7.6	17098	3	US-09-949-016-13119	Sequence 13119, A
12	142.5	7.5	899	4	US-09-297-648-2686	Sequence 2686, Ap
13	141	7.4	4608	3	US-09-475-515-76	Sequence 76, Appl
14	141	7.4	4689	3	US-09-475-515-74	Sequence 74, Appl
c 15	141	7.4	15543	3	US-09-949-016-17225	Sequence 17225, A
16	138.5	7.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
17	138.5	7.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
18	137	7.2	1042	3	US-09-533-559-5292	Sequence 5292, Ap
19	136.5	7.2	3671	3	US-09-382-552-232	Sequence 232, App
20	136.5	7.2	5915	3	US-09-382-552-3	Sequence 3, Appli
21	136.5	7.2	6240	3	US-09-382-552-117	Sequence 117, App
22	136.5	7.2	6909	3	US-09-382-552-21	Sequence 21, Appl
23	136.5	7.2	6910	3	US-09-382-552-15	Sequence 15, Appl
24	136.5	7.2	6911	3	US-09-382-552-1	Sequence 1, Appli
25	136.5	7.2	6911	3	US-09-382-552-14	Sequence 14, Appl
26	136.5	7.2	6911	3	US-09-382-552-16	Sequence 16, Appl
27	136.5	7.2	6911	3	US-09-382-552-17	Sequence 17, Appl
28	136.5	7.2	6911	3	US-09-382-552-18	Sequence 18, Appl
29	136.5	7.2	6911	3	US-09-382-552-19	Sequence 19, Appl
30	136.5	7.2	6911	3	US-09-382-552-20	Sequence 20, Appl
31	136.5	7.2	6912	3	US-09-382-552-13	Sequence 13, Appl
c 32	136	7.1	8280	5	US-09-936-271C-65	Sequence 65, Appl
c 33	136	7.1	10080	5	US-09-936-271C-97	Sequence 97, Appl
c 34	136	7.1	14077	3	US-09-902-540-1109	Sequence 1109, Ap
35	135.5	7.1	25431	3	US-09-949-016-13234	Sequence 13234, A
c 36	135.5	7.1	134008	3	US-09-949-016-13841	Sequence 13841, A
37	135	7.1	5877	3	US-10-152-886-54	Sequence 54, Appl
c 38	134.5	7.0	999	3	US-09-902-540-2460	Sequence 2460, Ap
c 39	134	7.0	30656	3	US-09-949-016-14613	Sequence 14613, A
c 40	132	6.9	2214	3	US-08-864-038A-1	Sequence 1, Appli
c 41	132	6.9	3331	3	US-08-864-038A-2	Sequence 2, Appli
c 42	132	6.9	3331	3	US-08-864-038A-4	Sequence 4, Appli
43	131	6.9	2889	2	US-08-537-002A-4	Sequence 4, Appli
44	131	6.9	2889	3	US-08-863-010-4	Sequence 4, Appli
45	131	6.9	2889	3	US-09-024-429-4	Sequence 4, Appli

# ALIGNMENTS

## RESULT 1

US-10-094-749-823

; Sequence 823, Application US/10094749

; Patent No. 6979557

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

```
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 823
; LENGTH: 2703
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-823
```

Alignment Scores:

```
Pred. No.: 4.81e-140 Length: 2703
Score: 1901.00 Matches: 340
Percent Similarity: 99.7% Conservative: 1
Best Local Similarity: 99.4% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 4 Gaps: 0
```

US-10-729-895A-2 (1-342) x US-10-094-749-823 (1-2703)

```
Qy      1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
      |||
Db      145 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204

Qy      21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40
      |||
Db      205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264

Qy      41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60
      |||
Db      265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 324

Qy      61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80
      |||
Db      325 TGCCACCACCGCCGAGCCAAGCACGCGCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384

Qy      81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100
      |||
Db      385 CTGATCGCTTACCGAGAAGCCACAAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 444

Qy     101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120
      |||
Db     445 CCAAACTATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA 504

Qy     121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140
      |||
Db     505 CCATACAGTGCCCTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 564

Qy     141 GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160
      |||
Db     565 GGTGGCAGTCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC 624

Qy     161 LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180
      |||
Db     625 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 684

Qy     181 AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200
      |||
Db     685 GACCTACAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 744

Qy     201 AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220
      |||
Db     745 GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 804

Qy     221 GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240
```

```

Db      805  |||GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864
Qy      241  GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260
Db      865  |||GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924
Qy      261  HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280
Db      925  |||CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 984
Qy      281  LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300
Db      985  |||CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044
Qy      301  GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320
Db      1045 |||CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA 1104
Qy      321  CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340
Db      1105 |||TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1164
Qy      341  ProSer 342
Db      1165 |||CCCAGC 1170

```

# RESULT 2

```

US-10-094-749-344
; Sequence 344, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 344
; LENGTH: 2042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-344

```

## Alignment Scores:

Pred. No.:	1.86e-27	Length:	2042
Score:	469.50	Matches:	119
Percent Similarity:	39.8%	Conservative:	26
Best Local Similarity:	32.7%	Mismatches:	96
Query Match:	24.6%	Indels:	123
DB:	4	Gaps:	12

US-10-729-895A-2 (1-342) x US-10-094-749-344 (1-2042)

Qy 11 LysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHisCysCys 30  
 Db 1102 CGAGAGCTGTGCCAGGAGTGAACAACCGCCCTACCTCTGTGAGAGTGGTCACTGCTGC 1161

Qy 31 GlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuValTrpThr 50  
 Db 1162 GGGGAGACTGGCTGCTGCACCTACTACTATGAGCTCTGGTGGTCTGGCTGCTCTGGACT 1221

Qy 51 IleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHisArgLeu 70  
 Db 1222 GTCCTCATCCTCTTTAGCTGCTGTGCGCCTTCGCCACCGACGAGCTAAACTCAGGCTG 1281

Qy 71 GlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAla---HisAsn 89  
 Db 1282 CAACAACAGCAGCGGAGCGTGAATCAACTTGTGGCTATCATGGGGCATGCCATGGG 1341

Qy 90 TyrSerAlaLeuPro-----PheTyrPheArgPheLeuProAsnTyrLeuLeu 105  
 Db 1342 GCTGGTCTTTCCCTACCGGTTCACTGCTTGACCTTCGCTTCTCAGCACCTTCAAGCCC 1401

Qy 106 ProProTyrGluGluValValAsnArgProProThrProProProTyrSerAlaPhe 125  
 Db 1402 CCAGCTACGAGGATGTGGTTCACGCCCCAGGCACACCACCCCCCTTATACTGTG--- 1458

Qy 126 GlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProPro 145  
 Db 1458 ----- 1458

Qy 146 GlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSer 165  
 Db 1459 -----GCCCGAGCGCGCCCTTGACTGCTTCCAGT 1488

Qy 166 ArgSerSerThrArgProProSerIleAlaAspProAspProSerAspLeuProValAsp 185  
 Db 1489 GAACAAACCTGCTGTCTCTCTCA-----TCCAGCTGCCCTGCCAC 1530

Qy 186 ArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGlu 205  
 Db 1531 TTTGAAGGAACAAATGTGAAGGTGT----- 1557

Qy 206 LeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAsp 225  
 Db 1557 ----- 1557

Qy 226 AspSerSerGluHisGlyAlaProAspSerLysGluLysThrProGly----- 241  
 Db 1558 ---TCCTCCCACAGAGTGCCCCCTCATCAGGAGGTGAGCCCGGGCAGGGTGACC 1614

Qy 242 -----ArgHisArgArgPheThrGlyAspSerGlyIleGlu 253  
 Db 1615 CCTGCCTCCACACCCCCCTCTGCCGCTATCGCCATTTAACTGGCGACTCCGGTATTGAG 1674

Qy 254 ValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGluPheAsn----- 270  
 Db 1675 CTCTGCCCTTGCTGCTGCCCTCCGGTGAGGGTGAGCCAGTCAAGGAGGTGAGGGTTAGTGCC 1734

Qy 271 ThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHisValArg 290  
 Db 1735 ACCCTGCCAGAT-----CTGGAGGACTACTCCCGGTGTGCACTACCC 1776

Qy 291 Pro-----ProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAla 307  
 Db 1777 CCAGAGTCTGTACCGCAGATCTTTCCCATGGGGCTGTCTTCCAGT----- 1821

Qy 308 ArgGluProGlyHisPro-----HisLeuProArgPro 318  
 Db 1822 ---GAAGGGGACATCCCATAGTAGTTTGTAGAGGGTGGATGGGTACTTGCCC----- 1872

Qy 319 ProAlaCysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSer 338  
 Db 1873 -----ACCAGAAACAGCCCTAGTCCCAACTCCTTGCGTTCC 1908

Qy 339 SerSerProSer 342  
 Db 1909 TTTGGCCCTCC 1920

## RESULT 3

US-09-513-999C-274

; Sequence 274, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6783961

; FILE REFERENCE: 59.US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C

; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 274

; LENGTH: 399

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 81..398

; FEATURE:

; NAME/KEY: sig\_peptide

; LOCATION: 81..152

; OTHER INFORMATION: score 10

; OTHER INFORMATION: seq LLLQLPSPSLA/RA

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 251

; OTHER INFORMATION: y=c or t

US-09-513-999C-274

## Alignment Scores:

Pred. No.:	7.43e-28	Length:	399
Score:	462.00	Matches:	77
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	24.2%	Indels:	0
DB:	3	Gaps:	0

US-10-729-895A-2 (1-342) x US-09-513-999C-274 (1-399)

```

Qy      9  GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      168 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 227

Qy      29  CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      228 TGCTGTGGACAGTCTCAGTGCTGYAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 287

Qy      49  TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      288 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 347

Qy      69  ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArg 85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      348 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGA 398

```

## RESULT 4

US-09-471-276-82

; Sequence 82, Application US/09471276

; Patent No. 6822072

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6822072

; FILE REFERENCE: GENSET.025CP1

; CURRENT APPLICATION NUMBER: US/09/471,276

; CURRENT FILING DATE: 1999-12-21

; EARLIER APPLICATION NUMBER: 09/057,719

; EARLIER FILING DATE: 1998-04-09

```
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCT/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pm
; SEQ ID NO 82
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 81..398
; NAME/KEY: sig_peptide
; LOCATION: 81..152
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10
; OTHER INFORMATION: seq LLLQALPSPLSA/RA
US-09-471-276-82
```

```
Alignment Scores:
Pred. No.:      7.43e-28      Length:      399
Score:          462.00        Matches:      77
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%    Mismatches:   0
Query Match:    24.2%          Indels:       0
DB:             3              Gaps:         0
```

US-10-729-895A-2 (1-342) x US-09-471-276-82 (1-399)

```
Qy      9 GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      168 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 227

Qy      29 CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      228 TGCTGTGGACAGTCTCAGTGCTGYAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 287

Qy      49 TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      288 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 347

Qy      69 ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArg 85
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      348 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGA 398
```

#### RESULT 5

```
US-09-247-155-182
; Sequence 182, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 182
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-247-155-182
```

```
Alignment Scores:
Pred. No.:      2.34e-26      Length:      1292
```

US-10-729-895A-2 (1-342) x US-09-247-155-182 (1-1292)

## RESULT 6

US-09-903-190-182

; Sequence 182. Application US/09903190

; Patent No. 6936692

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

; APPLICANT: Duclert, Aymeric

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: Complementary DNAs

; FILE REFERENCE: GENSET.021A

; CURRENT APPLICATION NUMBER: US/09/903,190

;  
; CURRENT FILING DATE: 2001-07-11



```
; PRIOR APPLICATION NUMBER: US/09/247,155A
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 182
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-903-190-182
```

```
Alignment Scores:
Pred. No.:      2.34e-26      Length:      1292
Score:          452.00        Matches:     101
Percent Similarity: 47.1%      Conservative: 23
Best Local Similarity: 38.4%    Mismatches:  77
Query Match:    23.7%         Indels:      62
DB:             3             Gaps:        7
```

US-10-729-895A-2 (1-342) x US-09-903-190-182 (1-1292)

```
Qy      11 LysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHisCysCys 30
      :::||| ||| ||| ||| ||| |||:::||||:||||:||||:||||:
Db      330 CGAGAGCTGTGCCAGGAGTGAATACCCAGCCCTACCTCTGTGAGACTGGTCATTGCTGT 389

Qy      31 GlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuValTrpThr 50
      |||:::|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      390 GGGGAGACTGGCTGCTGCACCTACTACTATGAACCTCTGGTGGTTCTGGCTGCTTTGGACT 449

Qy      51 IleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHisArgLeu 70
      :::::|:|:|: ||||| ||||| ||||| ||||| |||||
Db      450 GTCCTCATCTCTTTAGCTGCTGTGTGCTTCCGCCACCGAAGGGCTAAACTCAGGCTG 509

Qy      71 GlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAla---HisAsn 89
      ||| ||||| ||||| ||||| |||||:::|:|:|: ||| |||
Db      510 CAACAGCAACAGCGGCAGCGTGAAATCAACTTGTGGCTTACCATGGGGCATGCCACGGG 569

Qy      90 TyrSerAlaLeuPro-----PheTyrPheArgPheLeuProAsnTyrLeuLeu 105
      ::||| ||| ||| ::|
Db      570 GCTGGCCCTGTCCAACCGGTTCACTGCTTGACCTTCGCCTCCTCAGCGCCTCAAACCC 629

Qy      106 ProProTyrGluGluValValAsnArgProProThrProProProProTyrSerAlaPhe 125
      ||| |||||:::|||||::: ||| ||||| ||||| |||||:::
Db      630 CCAGCCTACGAGGATGTGGTTCAACACCCAGGCACACCGCCACCTCCTTACACTGTG--- 686

Qy      126 GlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProPro 145
      |||
Db      687 -----GGCCCA 692

Qy      146 GlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSer 165
      ||| ||| |||::: ::| ||| ||||| |||
Db      693 GGCTACCCTTGACTACTTCCAGTGAATGCACCCGCTGCTCT-----TCCGAATCCAGC 746

Qy      166 ArgSerSerThrArgProProSerIleAlaAspProAspProSerAspLeuProValAsp 185
      |||:::
Db      747 TGCTCTGCC-----CAC 758

Qy      186 ArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGlu 205
      ||| |||::: ||| |||
Db      759 TTGGAGGGGACAAATGTAGAAGGTGTTTCTCCAGCAGAGTGCTCTCCCTCACCAGGAG 818

Qy      206 LeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAsp 225
      ::||| ||| |||
Db      819 GGTGAGCCCAGGGCAGGATTG----- 839

Qy      226 AspSerSerGluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArg 245
      ||| ||| ||| |||:::|||||
Db      840 ---AGCCAGTTACATACCCCTTCCTGC-----CGCTATCGTCGC 878
```



```

Db      267 -----GCC 269
Qy      157 GlnSerSerProLeuSerGluProSerArgSerSerThrArgProProSerIleAlaAsp 176
          |||||::: ||| ::: |||
Db      270 CCAGGCCGCCCTTGACTGCTTCCAGTGAACAAACCTGCTGTCCTCCTCA----- 320
Qy      177 ProAspProSerAspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluPro 196
          ||| ||| ||| |||:::
Db      321 -----TCCAGCTGCCCTGCCCACTTTGAAGGAACAAATGTGAAGGTGTT----- 365
Qy      197 SerGlySerValAlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAla 216
Db      365 ----- 365
Qy      217 GluCysArgGluGluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLys 236
          ||||| ||||| :::
Db      366 -----TCCTCCCACCAGAGTGCCCCCCTCATCAG 395
Qy      237 GluLysThrProGly-----ArgHisArg 244
          ||| ||||| |||:::|||
Db      396 GAGGGTGAGCCCGGGCAGGGGTGACCCCTGCCTCCACACCCCTCCTGCCGCTATCGC 455
Qy      245 ArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAsp 264
          ||| |||||:::||| ||| :::
Db      456 CGTTTAACTGGCGACTCCGGTATTGAGCTCTGCCCTTGCTCCTGCCTCCGGTGAGGGTGAG 515
Qy      265 AspLeuLysGluPheAsn-----ThrLeuIleAspAspAlaLeuAspGlyProLeu 281
          :::||| ||||| ||| |||
Db      516 CCAGTCAAGGAGGTGAGGGTTAGTGCCACCCCTGCCAGAT-----CTG 557
Qy      282 AspPheCysAspSerCysHisValArgPro-----ProGlyAspGluGluGluGly 298
          ::: ||| ::: ||| ||| |||
Db      558 GAGGACTACTCCCCGTGTGCACTACCCCCAGAGTCTGTACCGCAGATCTTTCCCATGGGG 617
Qy      299 LeuCysGlnSerSerGluGluGlnAlaArgGluProGlyHisPro----- 313
          ||| ||| ||| |||
Db      618 CTGTCTTCCAGT-----GAAGGGGACATCCCATAGTAGTTTGA 659
Qy      314 -----HisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsnGlu 329
          :::||| |||
Db      660 GGGTGGATGGGTTACTTGCCC-----ACCAGAAAC 689
Qy      330 GlnAspSerProAsnSerGlnSerSerSerProSer 342
          ||||| ||| |||||
Db      690 AGCCCTAGTCCCACTCCTTGCGTTCCTTTGCCCTCC 728

```

RESULT 8

US-09-903-190-83

; Sequence 83, Application US/09903190

; Patent No. 6936692

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

; APPLICANT: Duclert, Aymeric

; APPLICANT: Bougueleret, Lydie

; TITLE OF INVENTION: Complementary DNAs

; FILE REFERENCE: GENSET.021A

; CURRENT APPLICATION NUMBER: US/09/903,190

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: US/09/247,155A

; PRIOR FILING DATE: 1999-02-09

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/074,121

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-09

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/081,563

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-13

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/096,116

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-10

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/099,273

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-04

; NUMBER OF SEQ ID NOS: 182

; SOFTWARE: Patent.pm

; SEQ ID NO 83

; LENGTH: 861

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: 612..644
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 829..834
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 850..861
US-09-903-190-83
```

Alignment Scores:

```
Pred. No.:      5.93e-17      Length:      861
Score:          329.50       Matches:     100
Percent Similarity: 36.0%    Conservative: 20
Best Local Similarity: 30.0% Mismatches:    90
Query Match:     17.3%      Indels:      123
DB:              3          Gaps:         12
```

US-10-729-895A-2 (1-342) x US-09-903-190-83 (1-861)

```
Qy      42 LeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysValCys 61
         |||||
Db      3 CTCTGGTGGTTCTGGCTGCTGGACTGTCTCATCTCTTTAGCTGCTGTTGCGCCTTC 62

Qy      62 HisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeu 81
         |||||
Db      63 CGCCACCGACGAGCTAAACTCAGGCTGCAACAACAGCAGCGGCAGCGTGAAATCAACTTG 122

Qy      82 IleAlaTyrArgGluAla---HisAsnTyrSerAlaLeuPro-----PheTyr 96
         ::||| ||| ||| |||
Db     123 TTGGCCTATCATGGGGCATGCCATGGGGCTGGTCTTTCCTACCGGTTCACTGCTTGAC 182

Qy      97 PheArgPheLeuProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProPro 116
         ||| ||| :: ||| |||||::|||::|||
Db     183 CTTCGCCTCCTCAGCACCTTCAAGCCCCAGCCTACGAGGATGTGGTTCACCGCCAGGC 242

Qy     117 ThrProProProProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGln 136
         |||||
Db     243 ACACCACCCCCCTTATACTGTG----- 266

Qy     137 CysGlyProAlaGlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAla 156
         |||
Db     267 -----GCC 269

Qy     157 GlnSerSerProLeuSerGluProSerArgSerSerThrArgProProSerIleAlaAsp 176
         |||||:: ||| :: |||
Db     270 CCAGGCCGCCCTTGACTGCTTCCAGTGAACAAACCTGCTGTCCTCCTCA----- 320

Qy     177 ProAspProSerAspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluPro 196
         ||| ||| ||| |||::
Db     321 -----TCCAGCTGCCCTGCCCACTTTGAAGGAACAAATGTGGAAGGTGTT----- 365

Qy     197 SerGlySerValAlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAla 216
         |||
Db     365 ----- 365

Qy     217 GluCysArgGluGluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLys 236
         ||||| ||||| ::
Db     366 -----TCCTCCCACCAGAGTGCCCCCCTCATCAG 395

Qy     237 GluLysThrProGly-----ArgHisArg 244
         ||| ||||| |||::|||
Db     396 GAGGGTGAGCCCGGGCAGGGTGACCCCTGCCTCCACCCCCCTCCTGCCGCTATCGC 455

Qy     245 ArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAsp 264
         ||| |||||
Db     456 CGTTTAAGTGGCGACTCCGGTATTGAGCTCTGCCCTGTCTCCTCGGTGAGGGTGAG 515

Qy     265 AspLeuLysGluPheAsn-----ThrLeuIleAspAspAlaLeuAspGlyProLeu 281
         ::||| ||| ||| |||
Db     516 CCAGTCAAGGAGGTGAGGGTAGTGCCACCCTGCCAGAT-----CTG 557

Qy     282 AspPheCysAspSerCysHisValArgPro-----ProGlyAspGluGluGluGly 298
         :: ||| :: ||| ||| |||
Db     558 GAGGACTACTCCCCGTGTGCACTACCCCCAGAGTCTGTACCGCAGATCTTCCCATGGGG 617
```

```

Qy      299 LeuCysGlnSerSerGluGluGlnAlaArgGluProGlyHisPro----- 313
          |||      |||      |||      |||
Db      618 CTGTCTTCCAGT-----GAAGGGGACATCCCATAAGTAGTTTGTGAGA 659

Qy      314 -----HisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsnGlu 329
          :::| |||||
Db      660 GGGTGGATGGGTTACTTGCCC-----ACCAGAAAC 689

Qy      330 GlnAspSerProAsnSerGlnSerSerSerSerProSer 342
          ||||| ||||| ||| |||||
Db      690 AGCCCTAGTCCCAACTCCTTGCCTTGGCCCTCC 728

```

RESULT 9

```

US-09-513-999C-3746
; Sequence 3746, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3746
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 126..455
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 126..233
; OTHER INFORMATION: score 9.7
; OTHER INFORMATION: seq LWTVLILFSCCCA/FX
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 237
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 281
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 361
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 378
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 406
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 410
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 411
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 2

```

```

; OTHER INFORMATION: Xaa=Cys or Gly or Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 16
; OTHER INFORMATION: Xaa=His or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 43
; OTHER INFORMATION: Xaa=Ala or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 49
; OTHER INFORMATION: Xaa=Cys or Gly or Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 58
; OTHER INFORMATION: Xaa=Asn or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 60
; OTHER INFORMATION: Xaa= * or Glu or Lys or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 63
; OTHER INFORMATION: Xaa=His or Gln
US-09-513-999C-3746

```

Alignment Scores:

Pred. No.:	2.61e-11	Length:	489
Score:	253.50	Matches:	52
Percent Similarity:	62.0%	Conservative:	10
Best Local Similarity:	52.0%	Mismatches:	33

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006
---------------------------------

# SCORE Search Results Details for Application 10729895 and Search Result 20070112\_180215\_us-10-729-895a-2.rnpbm.

[Score Home](#)  
[Page](#)

[Retrieve Application](#)  
[List](#)

[SCORE System](#)  
[Overview](#)

[SCORE](#)  
[FAQ](#)

[Comments /](#)  
[Suggestions](#)

This page gives you Search Results detail for the Application 10729895 and Search Result 20070112\_180215\_us-10-729-895a-2.rnpbm.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2007, 05:31:04 ; Search time 1483 Seconds  
(without alignments)  
4250.544 Million cell updates/sec

Title: US-10-729-895A-2  
Perfect score: 1909  
Sequence: 1 MPFLLGLRQDKEACVGTNNQ.....LLNTINEQDSPNSQSSSSPS 342

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abss/ABSSWEB\_spool/US10729895/runat\_12012007\_180215\_7466/app\_query.fasta\_1  
-DB=Published\_Applications\_NA\_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abss05p  
-USER=US10729895\_CGN\_1\_1\_2530@runat\_12012007\_180215\_7466 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA\_Main:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*

11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*  
 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*  
 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1909	100.0	1080	12	US-10-729-895-1	Sequence 1, Appli
2	1909	100.0	2976	7	US-10-297-880-7	Sequence 7, Appli
3	1909	100.0	4122	12	US-10-729-895-16	Sequence 16, Appl
4	1901	99.6	2703	7	US-10-094-749-823	Sequence 823, App
5	1868	97.9	1140	12	US-10-729-895-3	Sequence 3, Appli
6	1863	97.6	1508	7	US-10-291-172-32	Sequence 32, Appl
7	1863	97.6	1508	8	US-10-221-278-32	Sequence 32, Appl
8	1863	97.6	1550	7	US-10-119-428-35	Sequence 35, Appl
9	1863	97.6	3095	3	US-09-814-353-20113	Sequence 20113, A
10	1805	94.6	1715	7	US-10-291-172-408	Sequence 408, App
11	1805	94.6	1715	8	US-10-221-278-408	Sequence 408, App
12	1294	67.8	69081	6	US-10-087-192-1192	Sequence 1192, Ap
c 13	918	48.1	505	3	US-09-864-761-8443	Sequence 8443, Ap
c 14	753	39.4	684973	3	US-09-263-959-1	Sequence 1, Appli
15	751	39.3	439	3	US-09-814-353-14563	Sequence 14563, A
16	709.5	37.2	566	4	US-09-925-065A-842189	Sequence 842189,
17	709.5	37.2	566	5	US-09-925-065A-842189	Sequence 842189,
c 18	640	33.5	352	3	US-09-864-761-25171	Sequence 25171, A
19	631	33.1	382	3	US-09-814-353-1832	Sequence 1832, Ap
20	631	33.1	382	3	US-09-814-353-8179	Sequence 8179, Ap
21	579	30.3	462	3	US-09-918-995-9876	Sequence 9876, Ap
22	554.5	29.0	600	10	US-10-972-079-79549	Sequence 79549, A
23	525.5	27.5	599	10	US-10-972-079-79548	Sequence 79548, A
24	476.5	25.0	1199	3	US-09-925-301-110	Sequence 110, App
25	474.5	24.9	2047	7	US-10-108-260A-107	Sequence 107, App
26	474	24.8	1158	3	US-09-822-849A-15	Sequence 15, Appl
27	469.5	24.6	2042	7	US-10-094-749-344	Sequence 344, App
28	462	24.2	399	10	US-10-926-683-82	Sequence 82, Appl
29	452	23.7	1292	3	US-09-903-190-182	Sequence 182, App
30	452	23.7	1292	10	US-10-930-331-182	Sequence 182, App
c 31	450	23.6	528	12	US-10-301-480-503460	Sequence 503460,
c 32	450	23.6	528	12	US-10-301-480-1116869	Sequence 1116869,
33	450	23.6	561	12	US-10-301-480-461407	Sequence 461407,
34	450	23.6	561	12	US-10-301-480-1074816	Sequence 1074816,
c 35	450	23.6	562	4	US-09-925-065A-125966	Sequence 125966,
c 36	450	23.6	562	5	US-09-925-065A-125966	Sequence 125966,
c 37	450	23.6	568	4	US-09-925-065A-442874	Sequence 442874,
c 38	450	23.6	568	5	US-09-925-065A-442874	Sequence 442874,
c 39	450	23.6	571	12	US-10-301-480-223349	Sequence 223349,
c 40	450	23.6	571	12	US-10-301-480-836758	Sequence 836758,
c 41	449	23.5	558	12	US-10-301-480-503459	Sequence 503459,
c 42	449	23.5	558	12	US-10-301-480-1116868	Sequence 1116868,
c 43	449	23.5	569	4	US-09-925-065A-442873	Sequence 442873,
c 44	449	23.5	569	5	US-09-925-065A-442873	Sequence 442873,
45	448	23.5	561	12	US-10-301-480-461406	Sequence 461406,

#### ALIGNMENTS

RESULT 1  
 US-10-729-895-1  
 ; Sequence 1, Application US/10729895  
 ; Publication No. US20060063156A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: UNIVERSITY OF NEW MEXICO  
 ; TITLE OF INVENTION: OUTCOME PREDICTION AND RISK CLASSIFICATION IN CHILDHOOD  
 ; TITLE OF INVENTION: LEUKEMIA  
 ; FILE REFERENCE: N12-038US/310.00050101  
 ; CURRENT APPLICATION NUMBER: US/10/729,895  
 ; CURRENT FILING DATE: 2003-12-05



```
; PRIOR APPLICATION NUMBER: 60/510,904
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/510,968
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/432,064
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 60/432,077
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 60/432,078
; PRIOR FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1026)
US-10-729-895-1
```

Alignment Scores:

```
Pred. No.:      1.11e-177      Length:      1080
Score:          1909.00      Matches:      342
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:            12          Gaps:        0
```

US-10-729-895A-2 (1-342) x US-10-729-895-1 (1-1080)

```
Qy      1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
      |||
Db      1 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60

Qy     21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40
      |||
Db     61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120

Qy     41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60
      |||
Db    121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180

Qy     61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80
      |||
Db    181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240

Qy     81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100
      |||
Db    241 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTG 300

Qy    101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120
      |||
Db    301 CCAAATATTATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA 360

Qy    121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140
      |||
Db    361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCA 420

Qy    141 GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160
      |||
Db    421 GGTGGCAGTCCCCGGGCATCGATCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC 480

Qy    161 LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180
      |||
Db    481 TTGTCTGAGCCAGCAGAAGCAGCACAAAGCCCCAAGCATCGCTGACCCTGATCCCTCT 540

Qy    181 AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200
      |||
Db    541 GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 600

Qy    201 AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220
      |||
Db    601 GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660

Qy    221 GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240
```

```

Db      661  |||||GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
Qy      241  GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260
Db      721  |||||GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 780
Qy      261  HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280
Db      781  |||||CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
Qy      281  LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300
Db      841  |||||CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 900
Qy      301  GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320
Db      901  |||||CACTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
Qy      321  CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340
Db      961  |||||TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCAACCTCCCAGAGCAGCAGCTCC 1020
Qy      341  ProSer 342
Db      1021 |||||CCCAGC 1026

```

# RESULT 2

US-10-297-880-7

```

; Sequence 7, Application US/10297880
; Publication No. US20030211513A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: HE, Ann; NGUYEN, Danniel B.
; APPLICANT: YAO, Monique G.; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; TANG, Y. Tom
; APPLICANT: XU, Yuming; HAFALIA, April J.A.
; APPLICANT: AZIMZAI, Yalda; CHAWLA, Narinder K.
; TITLE OF INVENTION: INTRACELLULAR SIGNALING PROTEINS
; FILE REFERENCE: PF-0782 USN
; CURRENT APPLICATION NUMBER: US/10/297,880
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US01/18595
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: US 60/210,582
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: US 60/212,443
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1478005CB1
US-10-297-880-7

```

# Alignment Scores:

Pred. No.:	3.92e-177	Length:	2976
Score:	1909.00	Matches:	342
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-729-895A-2 (1-342) x US-10-297-880-7 (1-2976)

```

Qy      1  MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
Db      479 |||||ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 538
Qy      21  SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40
Db      539 |||||AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 598

```

Qy 41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60  
 |||||  
 Db 599 GAACTCTGGTGGTTCTGGCTGGTGGACCATCATCATCTCTGAGCTGCTGTGTGTT 658

Qy 61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80  
 |||||  
 Db 659 TGCCACCACCGCCGAGCCAAGCACCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 718

Qy 81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100  
 |||||  
 Db 719 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTG 778

Qy 101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120  
 |||||  
 Db 779 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 838

Qy 121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140  
 |||||  
 Db 839 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGCCCTGCA 898

Qy 141 GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160  
 |||||  
 Db 899 GGTGGCAGTCCCCGGGCATCGATCCACACAGGGGATCCAGGGGGACAGAGCAGCCCC 958

Qy 161 LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180  
 |||||  
 Db 959 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCTCT 1018

Qy 181 AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200  
 |||||  
 Db 1019 GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 1078

Qy 201 AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220  
 |||||  
 Db 1079 GCTGGCTGGGGAGCTGGACCCGGGGGCTTCTGGACAAAGATGCAGAATGTAGGGAG 1138

Qy 221 GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240  
 |||||  
 Db 1139 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 1198

Qy 241 GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260  
 |||||  
 Db 1199 GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGCAACCGGGGC 1258

Qy 261 HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280  
 |||||  
 Db 1259 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 1318

Qy 281 LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300  
 |||||  
 Db 1319 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1378

Qy 301 GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320  
 |||||  
 Db 1379 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCGCA 1438

Qy 321 CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340  
 |||||  
 Db 1439 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC 1498

Qy 341 ProSer 342  
 |||||  
 Db 1499 CCCAGC 1504

RESULT 3

US-10-729-895-16

; Sequence 16, Application US/10729895

; Publication No. US20060063156A1

; GENERAL INFORMATION:

; APPLICANT: UNIVERSITY OF NEW MEXICO

; TITLE OF INVENTION: OUTCOME PREDICTION AND RISK CLASSIFICATION IN CHILDHOOD

; TITLE OF INVENTION: LEUKEMIA

; FILE REFERENCE: N12-038US/310.00050101

; CURRENT APPLICATION NUMBER: US/10/729,895

; CURRENT FILING DATE: 2003-12-05

; PRIOR APPLICATION NUMBER: 60/510,904  
 ; PRIOR FILING DATE: 2003-10-14  
 ; PRIOR APPLICATION NUMBER: 60/510,968  
 ; PRIOR FILING DATE: 2003-10-14  
 ; PRIOR APPLICATION NUMBER: 60/432,064  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: 60/432,077  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: 60/432,078  
 ; PRIOR FILING DATE: 2002-12-06  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 16  
 ; LENGTH: 4122  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-729-895-16

Alignment Scores:

Pred. No.:	5.87e-177	Length:	4122
Score:	1909.00	Matches:	342
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-729-895A-2 (1-342) x US-10-729-895-16 (1-4122)

Qy	1	MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln	20
Db	145	ATGCCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	204
Qy	21	SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr	40
Db	205	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	264
Qy	41	GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal	60
Db	265	GAACTCTGGTGGTCTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	324
Qy	61	CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn	80
Db	325	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	384
Qy	81	LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu	100
Db	385	CTGATCGCTTACCAGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTC	444
Qy	101	ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro	120
Db	445	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA	504
Qy	121	ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla	140
Db	505	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	564
Qy	141	GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro	160
Db	565	GGTGGCAGTCCCCGGGCATCGATCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCC	624
Qy	161	LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer	180
Db	625	TTGTCTGAGCCAGCAGAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	684
Qy	181	AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal	200
Db	685	GACCTACAGTTGACCGAGCAGCCACAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	744
Qy	201	AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu	220
Db	745	GCTGGCCTGGGGGAGCTGGACCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	804
Qy	221	GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro	240
Db	805	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	864

```

Qy      241 GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260
      |||
Db      865 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924

Qy      261 HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280
      |||
Db      925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC 984

Qy      281 LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300
      |||
Db      985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy      301 GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320
      |||
Db     1045 CAGTCTCTGAGGAGCAGGCTCGAGAGCTGGGCACCCGCACCTGCCACGCCGCCGCCGA 1104

Qy      321 CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340
      |||
Db     1105 TGCCTGCTGTGAACACCATCAACGAGCAGGACTCTCCCACTCCAGAGCAGCAGCTCC 1164

Qy      341 ProSer 342
      |||
Db     1165 CCCAGC 1170

```

#### RESULT 4

US-10-094-749-823

; Sequence 823, Application US/10094749

; Publication No. US20030219741A1

#### ; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHIKO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA

; FILE REFERENCE: 084335/0160

; CURRENT APPLICATION NUMBER: US/10/094,749

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435

; PRIOR FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14

; NUMBER OF SEQ ID NOS: 3381

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 823

; LENGTH: 2703

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-094-749-823

#### Alignment Scores:

Pred. No.:	2.13e-176	Length:	2703
Score:	1901.00	Matches:	340
Percent Similarity:	99.7%	Conservative:	1
Best Local Similarity:	99.4%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	7	Gaps:	0

US-10-729-895A-2 (1-342) x US-10-094-749-823 (1-2703)

```

Qy      1 MetProPheLeuLeuGlyLeuArgGlnAspLysGluAlaCysValGlyThrAsnAsnGln 20
      |||
Db     145 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204

```

Qy 21 SerTyrIleCysAspThrGlyHisCysCysGlyGlnSerGlnCysCysAsnTyrTyrTyr 40  
 |||  
 Db 205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264

Qy 41 GluLeuTrpTrpPheTrpLeuValTrpThrIleIleIleIleLeuSerCysCysCysVal 60  
 |||  
 Db 265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 324

Qy 61 CysHisHisArgArgAlaLysHisArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsn 80  
 |||  
 Db 325 TGCCACCACCGCCGAGCCAAGCACCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384

Qy 81 LeuIleAlaTyrArgGluAlaHisAsnTyrSerAlaLeuProPheTyrPheArgPheLeu 100  
 |||  
 Db 385 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTGT 444

Qy 101 ProAsnTyrLeuLeuProProTyrGluGluValValAsnArgProProThrProProPro 120  
 |||  
 Db 445 CCAAATATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA 504

Qy 121 ProTyrSerAlaPheGlnLeuGlnGlnGlnGlnLeuLeuProProGlnCysGlyProAla 140  
 |||  
 Db 505 CCATACAGTGCCTTCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCTGCA 564

Qy 141 GlyGlySerProProGlyIleAspProThrArgGlySerGlnGlyAlaGlnSerSerPro 160  
 |||  
 Db 565 GGTGGCAGTCCCCGGGCATCGATCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC 624

Qy 161 LeuSerGluProSerArgSerSerThrArgProProSerIleAlaAspProAspProSer 180  
 |||  
 Db 625 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCTCT 684

Qy 181 AspLeuProValAspArgAlaAlaThrLysAlaProGlyMetGluProSerGlySerVal 200  
 |||  
 Db 685 GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 744

Qy 201 AlaGlyLeuGlyGluLeuAspProGlyAlaPheLeuAspLysAspAlaGluCysArgGlu 220  
 |||  
 Db 745 GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAG 804

Qy 221 GluLeuLeuLysAspAspSerSerGluHisGlyAlaProAspSerLysGluLysThrPro 240  
 |||  
 Db 805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864

Qy 241 GlyArgHisArgArgPheThrGlyAspSerGlyIleGluValCysValCysAsnArgGly 260  
 |||  
 Db 865 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924

Qy 261 HisHisAspAspAspLeuLysGluPheAsnThrLeuIleAspAspAlaLeuAspGlyPro 280  
 |||  
 Db 925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 984

Qy 281 LeuAspPheCysAspSerCysHisValArgProProGlyAspGluGluGluGlyLeuCys 300  
 |||  
 Db 985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy 301 GlnSerSerGluGluGlnAlaArgGluProGlyHisProHisLeuProArgProProAla 320  
 |||  
 Db 1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGACCTGCCACGGCCGCCCTCA 1104

Qy 321 CysLeuLeuLeuAsnThrIleAsnGluGlnAspSerProAsnSerGlnSerSerSerSer 340  
 |||  
 Db 1105 TGCTGTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCAGAGCAGCAGCTCC 1164

Qy 341 ProSer 342  
 |||  
 Db 1165 CCCAGC 1170

RESULT 5

US-10-729-895-3

; Sequence 3, Application US/10729895

; Publication No. US20060063156A1

; GENERAL INFORMATION:

; APPLICANT: UNIVERSITY OF NEW MEXICO

; TITLE OF INVENTION: OUTCOME PREDICTION AND RISK CLASSIFICATION IN CHILDHOOD

```
; TITLE OF INVENTION: LEUKEMIA
; FILE REFERENCE: N12-038US/310.00050101
; CURRENT APPLICATION NUMBER: US/10/729,895
; CURRENT FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: 60/510,904
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/510,968
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/432,064
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 60/432,077
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 60/432,078
; PRIOR FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 3
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1089)
US-10-729-895-3
```

Alignment Scores:

```
Pred. No.:      1.3e-173      Length:      1140
Score:          1868.00      Matches:      334
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    97.9%      Indels:      0
DB:             12          Gaps:         0
```

US-10-729-895A-2 (1-342) x US-10-729-895-3 (1-1140)

```
Qy      9  GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28
|||||
Db      88  CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 147

Qy      29  CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48
|||||
Db     148  TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAAGCTGGTGGTTCTGGCTGGTG 207

Qy      49  TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68
|||||
Db     208  TGGACCATCATCATCATCCTGAGTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC 267

Qy      69  ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis 88
|||||
Db     268  CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 327

Qy      89  AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr 108
|||||
Db     328  AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTCGCAAATATTACTACCTCCTTAT 387

Qy     109  GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln 128
|||||
Db     388  GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCTTCCAGCTACAG 447

Qy     129  GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp 148
|||||
Db     448  CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT 507

Qy     149  ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer 168
|||||
Db     508  CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCCTTGTCTGAGCCAGCAGAAGCAGC 567

Qy     169  ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla 188
|||||
Db     568  ACAAGACCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 627

Qy     189  ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro 208
|||||
Db     628  ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 687

Qy     209  GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer 228
```

```

Db      688  ||| GGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 747
Qy      229  GluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGly 248
Db      748  ||| GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 807
Qy      249  AspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGlu 268
Db      808  ||| GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 867
Qy      269  PheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHis 288
Db      868  ||| TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 927
Qy      289  ValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArg 308
Db      928  ||| GTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGTCTAGTCTCTGAGGAGCAGGCTCGA 987
Qy      309  GluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsn 328
Db      988  ||| GAGCCTGGGCACCCGCACCTGCCACGGCCGCCGCATGCCTGCTGCTGAACACCATCAAC 1047
Qy      329  GluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342
Db      1048 ||| GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCCAGC 1089

```

# RESULT 6

```

US-10-291-172-32
; Sequence 32, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 32
; LENGTH: 1508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(1175)
US-10-291-172-32

```

## Alignment Scores:

Pred. No.:	5.72e-173	Length:	1508
Score:	1863.00	Matches:	333
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	97.6%	Indels:	0
DB:	7	Gaps:	0

US-10-729-895A-2 (1-342) x US-10-291-172-32 (1-1508)

```

Qy      9  GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28
Db      174 ||| CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 233
Qy      29  CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48
Db      234 ||| TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAAGTCTGGTGGTTCTGGCTGGTG 293

```



Qy 49 TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68  
 |||  
 Db 294 TGGACCATCATCATCATCTGAGTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC 353

Qy 69 ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis 88  
 |||  
 Db 354 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 413

Qy 89 AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr 108  
 |||  
 Db 414 AATTACTCAGCGCTGCCATTTTATTTCAGGTTTTTGCCAACTATTTACTACCTCCTTAT 473

Qy 109 GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln 128  
 |||  
 Db 474 GAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATAACAGTGCTTCCAGCTACAG 533

Qy 129 GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp 148  
 |||  
 Db 534 CAGCAGCAGCTGCTGCCTCCACAGTGTGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 593

Qy 149 ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer 168  
 |||  
 Db 594 CCCACCAGGGGATCCAGGGGGCAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC 653

Qy 169 ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla 188  
 |||  
 Db 654 ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 713

Qy 189 ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro 208  
 |||  
 Db 714 ACCAAGCCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCTGGGGGAGCTGGACCCG 773

Qy 209 GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer 228  
 |||  
 Db 774 GGGGCTTCTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 833

Qy 229 GluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGly 248  
 |||  
 Db 834 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTACAGGT 893

Qy 249 AspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGlu 268  
 |||  
 Db 894 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 953

Qy 269 PheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHis 288  
 |||  
 Db 954 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT 1013

Qy 289 ValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArg 308  
 |||  
 Db 1014 GTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGTCAGCCCTCTGAGGAGCAGGCTCGA 1073

Qy 309 GluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsn 328  
 |||  
 Db 1074 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCGCATGCCTGCTGTAACACCATCAAC 1133

Qy 329 GluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342  
 |||  
 Db 1134 GAGCAGGACTCTCCCAACTCCAGAGCAGCAGCTCCCCCAGC 1175

# RESULT 7

US-10-221-278-32

; Sequence 32, Application US/10221278

; Publication No. US20040034208A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: No. US20040034208A1e1 Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-045

; CURRENT APPLICATION NUMBER: US/10/221,278

; CURRENT FILING DATE: 2002-09-06

; PRIOR APPLICATION NUMBER: 09/693,267

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 09/665,363

; PRIOR FILING DATE: 2000-09-19

; PRIOR APPLICATION NUMBER: 09/616,847

; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 09/596,193  
 ; PRIOR FILING DATE: 2000-06-17  
 ; PRIOR APPLICATION NUMBER: 09/574,454  
 ; PRIOR FILING DATE: 2000-05-19  
 ; PRIOR APPLICATION NUMBER: 09/519,705  
 ; PRIOR FILING DATE: 2000-03-07  
 ; NUMBER OF SEQ ID NOS: 752  
 ; SEQ ID NO 32  
 ; LENGTH: 1508  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (87)..(1175)  
 US-10-221-278-32

Alignment Scores:

Pred. No.:	5.72e-173	Length:	1508
Score:	1863.00	Matches:	333
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	97.6%	Indels:	0
DB:	8	Gaps:	0

US-10-729-895A-2 (1-342) x US-10-221-278-32 (1-1508)

QY	9	GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis	28
Db	174	CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC	233
QY	29	CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal	48
Db	234	TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG	293
QY	49	TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis	68
Db	294	TGGACCATCATCATCATCCTGAGTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC	353
QY	69	ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis	88
Db	354	CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC	413
QY	89	AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr	108
Db	414	AATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTGCCAACTATTTACTACCTCCTTAT	473
QY	109	GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln	128
Db	474	GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCTTCCAGCTACAG	533
QY	129	GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp	148
Db	534	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT	593
QY	149	ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer	168
Db	594	CCCACCAGGGGATCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC	653
QY	169	ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla	188
Db	654	ACAAGACCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC	713
QY	189	ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro	208
Db	714	ACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG	773
QY	209	GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer	228
Db	774	GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT	833
QY	229	GluHisGlyAlaProAspSerLysGluLysThrProGlyArgHisArgArgPheThrGly	248
Db	834	GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT	893
QY	249	AspSerGlyIleGluValCysValCysAsnArgGlyHisHisAspAspAspLeuLysGlu	268

```

Db      894  GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 953
Qy      269  PheAsnThrLeuIleAspAspAlaLeuAspGlyProLeuAspPheCysAspSerCysHis 288
Db      954  TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCAT 1013
Qy      289  ValArgProProGlyAspGluGluGluGlyLeuCysGlnSerSerGluGluGlnAlaArg 308
Db      1014 GTGCGGCCCCCTGGTGTATGAGGAGGAAGGCCTCTGTGAGCCCTCTGAGGAGCAGGCTCGA 1073
Qy      309  GluProGlyHisProHisLeuProArgProProAlaCysLeuLeuLeuAsnThrIleAsn 328
Db      1074 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1133
Qy      329  GluGlnAspSerProAsnSerGlnSerSerSerSerProSer 342
Db      1134 GAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCCCCCAGC 1175

```

# RESULT 8

```

US-10-119-428-35
; Sequence 35, Application US/10119428
; Publication No. US20030165881A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Wehrman, Tom
; APPLICANT: Ren, Feiyan
; APPLICANT: Ma, Yunqing
; APPLICANT: Zhou, Ping
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030165881A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 789CIP2
; CURRENT APPLICATION NUMBER: US/10/119,428
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 35
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (129)..(1220)
US-10-119-428-35

```

## Alignment Scores:

Pred. No.:	5.92e-173	Length:	1550
Score:	1863.00	Matches:	333
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	97.6%	Indels:	0
DB:	7	Gaps:	0

US-10-729-895A-2 (1-342) x US-10-119-428-35 (1-1550)

```

Qy      9  GlnAspLysGluAlaCysValGlyThrAsnAsnGlnSerTyrIleCysAspThrGlyHis 28
Db      216 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 275
Qy      29  CysCysGlyGlnSerGlnCysCysAsnTyrTyrTyrGluLeuTrpTrpPheTrpLeuVal 48
Db      276 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 335
Qy      49  TrpThrIleIleIleIleLeuSerCysCysCysValCysHisHisArgArgAlaLysHis 68

```

```

Db      336 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCGAGCCAAGCAC 395
Qy      69 ArgLeuGlnAlaGlnGlnArgGlnHisGluIleAsnLeuIleAlaTyrArgGluAlaHis 88
Db      396 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 455
Qy      89 AsnTyrSerAlaLeuProPheTyrPheArgPheLeuProAsnTyrLeuLeuProProTyr 108
Db      456 AATTACTCAGCGCTGCCATTTTATTTCAGGTTTTGCCAACTATTACTACCTCCTTAT 515
Qy      109 GluGluValValAsnArgProProThrProProProProTyrSerAlaPheGlnLeuGln 128
Db      516 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATAACAGTGCCTTCCAGCTACAG 575
Qy      129 GlnGlnGlnLeuLeuProProGlnCysGlyProAlaGlyGlySerProProGlyIleAsp 148
Db      576 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 635
Qy      149 ProThrArgGlySerGlnGlyAlaGlnSerSerProLeuSerGluProSerArgSerSer 168
Db      636 CCCACCAGGGGATCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC 695
Qy      169 ThrArgProProSerIleAlaAspProAspProSerAspLeuProValAspArgAlaAla 188
Db      696 ACAAGACCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 755
Qy      189 ThrLysAlaProGlyMetGluProSerGlySerValAlaGlyLeuGlyGluLeuAspPro 208
Db      756 ACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 815
Qy      209 GlyAlaPheLeuAspLysAspAlaGluCysArgGluGluLeuLeuLysAspAspSerSer 228

```

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006
---------------------------------



7	82	7.6	489	3	US-09-513-999C-3746	Sequence 3746, Ap
8	80.4	7.4	861	3	US-09-247-155-83	Sequence 83, Appl
9	80.4	7.4	861	3	US-09-903-190-83	Sequence 83, Appl
10	43.2	4.0	1903	3	US-09-148-545-55	Sequence 55, Appl
11	43.2	4.0	1903	3	US-09-621-011-55	Sequence 55, Appl
12	40.6	3.8	519	3	US-09-252-991A-3601	Sequence 3601, Ap
13	40.6	3.8	939	3	US-09-252-991A-3645	Sequence 3645, Ap
14	40.6	3.8	1881	3	US-09-252-991A-3697	Sequence 3697, Ap
c 15	40.6	3.8	1998	3	US-09-252-991A-3812	Sequence 3812, Ap
c 16	39.4	3.6	719	3	US-09-620-312D-669	Sequence 669, App
c 17	38.8	3.6	978	3	US-09-252-991A-86	Sequence 86, Appl
c 18	38.8	3.6	1206	3	US-09-252-991A-92	Sequence 92, Appl
c 19	38	3.5	2199	3	US-09-902-540-6388	Sequence 6388, Ap
c 20	38	3.5	2585	3	US-09-902-540-424	Sequence 424, App
c 21	37.6	3.5	480	3	US-09-252-991A-2705	Sequence 2705, Ap
c 22	37.6	3.5	594	3	US-09-252-991A-2890	Sequence 2890, Ap
23	37.6	3.5	784	3	US-10-017-754-1866	Sequence 1866, Ap
24	37.6	3.5	785	3	US-10-017-754-1868	Sequence 1868, Ap
25	37.6	3.5	789	3	US-10-017-754-1867	Sequence 1867, Ap
26	37.6	3.5	790	3	US-10-017-754-1865	Sequence 1865, Ap
27	37.6	3.5	1635	3	US-09-234-332-4	Sequence 4, Appli
28	37.6	3.5	1635	3	US-09-702-705-1798	Sequence 1798, Ap
29	37.6	3.5	1635	3	US-09-736-457-1798	Sequence 1798, Ap
30	37.6	3.5	1635	3	US-09-671-325-1798	Sequence 1798, Ap
31	37.6	3.5	1635	3	US-10-017-754-1798	Sequence 1798, Ap
c 32	37.6	3.5	23901	3	US-09-949-016-16773	Sequence 16773, A
c 33	37.4	3.5	7218	2	US-08-232-463-14	Sequence 14, Appl
c 34	37.2	3.4	505	3	US-09-621-976-15639	Sequence 15639, A
35	37.2	3.4	1500	2	US-08-487-037-4	Sequence 4, Appli
36	37.2	3.4	5204	3	US-10-037-417-7	Sequence 7, Appli
37	37.2	3.4	5640	3	US-09-620-312D-41	Sequence 41, Appl
38	37.2	3.4	8296	3	US-10-037-182-35	Sequence 35, Appl
39	37.2	3.4	11350	3	US-10-037-182-1	Sequence 1, Appli
40	37	3.4	5877	3	US-10-152-886-54	Sequence 54, Appl
c 41	36.8	3.4	12055	3	US-09-949-016-13233	Sequence 13233, A
c 42	36.6	3.4	465	3	US-09-902-540-7880	Sequence 7880, Ap
43	36.6	3.4	639	3	US-09-252-991A-13260	Sequence 13260, A
c 44	36.6	3.4	942	4	US-09-754-949-1	Sequence 1, Appli
c 45	36.6	3.4	942	5	US-10-417-422-1	Sequence 1, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-094-749-823

; Sequence 823, Application US/10094749

; Patent No. 6979557

#### GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHIKO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA

; FILE REFERENCE: 084335/0160

; CURRENT APPLICATION NUMBER: US/10/094,749

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435

; PRIOR FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14

; NUMBER OF SEQ ID NOS: 3381

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 823  
; LENGTH: 2703  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-094-749-823

Query Match 99.7%; Score 1076.8; DB 4; Length 2703;  
Best Local Similarity 99.8%; Pred. No. 1.1e-277;  
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60
          |||
Db      145 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204

Qy      61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120
          |||
Db      205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264

Qy      121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCATCCTGAGCTGCTGCTGTGTT 180
          |||
Db      265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCATCCTGAGCTGCTGCTGTGTT 324

Qy      181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240
          |||
Db      325 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384

Qy      241 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 300
          |||
Db      385 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 444

Qy      301 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 360
          |||
Db      445 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 504

Qy      361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 420
          |||
Db      505 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 564

Qy      421 GGTGGCAGTCCCCCGGGCATCGATCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCC 480
          |||
Db      565 GGTGGCAGTCCCCCGGGCATCGATCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCC 624

Qy      481 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540
          |||
Db      625 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 684

Qy      541 GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAAGTGGCTCTGTG 600
          |||
Db      685 GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAAGTGGCTCTGTG 744

Qy      601 GCTGGCCTGGGGAGCTGGACCCGGGGGCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
          |||
Db      745 GCTGGCCTGGGGAGCTGGACCCGGGGGCTTCTGGACAAAGATGCAGAATGTAGGGAG 804

Qy      661 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
          |||
Db      805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864

Qy      721 GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG 780
          |||
Db      865 GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG 924

Qy      781 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
          |||
Db      925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 984

Qy      841 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT 900
          |||
Db      985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy      901 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGACCTGCCACGGCCGCCGCA 960
          |||
Db      1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGACCTGCCACGGCCGCCCTCA 1104

Qy      961 TGCTGTCTGTAACACCATCAACGAGCAGGACTCTCCAACTCCAGAGCAGCAGCTCC 1020
```

```

|||||
Db      1105 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1164
QY      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
|||||
Db      1165 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1224

```

# RESULT 2

US-09-513-999C-274

; Sequence 274, Application US/09513999C

; Patent No. 6783961

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6783961

; FILE REFERENCE: 59.US2.REG

; CURRENT APPLICATION NUMBER: US/09/513,999C

; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 274

; LENGTH: 399

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 81..398

; FEATURE:

; NAME/KEY: sig\_peptide

; LOCATION: 81..152

; OTHER INFORMATION: score 10

; OTHER INFORMATION: seq LLLLQALPSPLSA/RA

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 251

; OTHER INFORMATION: y=c or t

US-09-513-999C-274

Query Match 21.4%; Score 231.6; DB 3; Length 399;

Best Local Similarity 99.6%; Pred. No. 4.6e-52;

Matches 231; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84
|||||
Db      168 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 227

QY      85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
|||||
Db      228 TGCTGTGGACAGTCTCAGTGCTGYAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 287

QY      145 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 204
|||||
Db      288 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 347

QY      205 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAG 256
|||||
Db      348 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAG 399

```

# RESULT 3

US-09-471-276-82

; Sequence 82, Application US/09471276

; Patent No. 6822072

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6822072

; FILE REFERENCE: GENSET.025CP1

; CURRENT APPLICATION NUMBER: US/09/471,276



```
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCT/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pm
; SEQ ID NO 82
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 81..398
; NAME/KEY: sig_peptide
; LOCATION: 81..152
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10
; OTHER INFORMATION: seq LLLQALPSPLSA/RA
US-09-471-276-82
```

```
Query Match          21.4%; Score 231.6; DB 3; Length 399;
Best Local Similarity 99.6%; Pred. No. 4.6e-52;
Matches 231; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      168 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 227

Qy      85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      228 TGCTGTGGACAGTCTCAGTGCTGYAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 287

Qy      145 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGGCCACCACCGCCGAGCCAAGCAC 204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      288 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGGCCACCACCGCCGAGCCAAGCAC 347

Qy      205 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAG 256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      348 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAG 399
```

```
RESULT 4
US-10-094-749-344
; Sequence 344, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
```

Query Match 11.4%; Score 123.2; DB 4; Length 2042;  
Best Local Similarity 73.1%; Pred. No. 8.5e-23;  
Matches 158; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

## RESULT 5

Query Match 11.4%; Score 122.6; DB 3; Length 1292;  
Best Local Similarity 74.2%; Pred. No. 1e-22;  
Matches 155; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

<http://es/ScoreAccessWeb/GetItem.action?AppId=10729895&seqId=1038738&ItemName...> 1/22/2007

US-09-903-190-182

US-09-903-190-182

Qy	52	AACAATCAAAGCTACACTCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAAC	111
Db	351	AATACCCAGCCCTACCTCTGTGAGACTGGTCATTGCTGTGGGGAGACTGGCTGCTGCACC	410
Qy	112	TACTACTATGAACCTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCATCTGAGCTGC	171
Db	411	TACTACTATGAACCTCTGGTGGTTCTGGCTGCTTTGGACTGTCTCTATCCTCTTTAGCTGC	470
Qy	172	TGCTGTGTTTGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACAT	231
Db	471	TGTTGTGCCTTCGCCACCGAAGGGCTAAACTCAGGCTGCAACAGCAACAGCGGCAGCGT	530
Qy	232	GAAATCAACCTGATCGCTTACCGAGAAGC	260
Db	531	GAAATCAACTTGTTGGCTTACCATGGGGC	559

US-09-513-999C-3746

; LOCATION: 126..455

```

; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 126..233
; OTHER INFORMATION: score 9.7
; OTHER INFORMATION: seq LWTVLILFSCCCA/FX
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 237
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 281
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 361
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 378
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 406
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 410
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 411
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 2
; OTHER INFORMATION: Xaa=Cys or Gly or Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 16
; OTHER INFORMATION: Xaa=His or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 43
; OTHER INFORMATION: Xaa=Ala or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 49
; OTHER INFORMATION: Xaa=Cys or Gly or Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 58
; OTHER INFORMATION: Xaa=Asn or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 60
; OTHER INFORMATION: Xaa= * or Glu or Lys or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 63
; OTHER INFORMATION: Xaa=His or Gln
US-09-513-999C-3746

```

```

Query Match          7.6%; Score 82; DB 3; Length 489;
Best Local Similarity 72.7%; Pred. No. 4.9e-12;
Matches 117; Conservative 1; Mismatches 42; Indels 1; Gaps 1;

```

```

Qy      107 GCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCTCTGA 166
          ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      160 GCACCTACTACTATGAGCTCTGGTGGTTCTGGCTGCTCTGGACTGTCTCATCTCTTTA 219
Qy      167 GCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGC 226

```

```

Db      220 GCTGCTGTTGCGCCTTCNGCCACCGACGAGCTAAACTCAGGCTGCAACAACAGCAGCGGC 279
Qy      227 AACATGAAATCAACCTGATCGCTTACCGAGAAGCCCACAAT 267
      1: ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db      280 -ASGTGAAATCAACTTGTGGCCTATCATGGGGCATGCCAT 319

```

# RESULT 8

```

US-09-247-155-83
; Sequence 83, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 83
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 612..644
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 829..834
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 850..861
US-09-247-155-83

```

```

Query Match          7.4%; Score 80.4; DB 3; Length 861;
Best Local Similarity 71.9%; Pred. No. 1.7e-11;
Matches 105; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

```

Qy      122 AACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCATCTGAGCTGCTGCTGTGTTT 181
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 AGCTCTGGTGGTTCTGGCTGCTCTGGACTGTCCTCATCCTCTTAGCTGCTGTTGCGCCT 60

Qy      182 GCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACC 241
      || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 TCCGCCACCGACGAGCTAAACTCAGGCTGCAACAACAGCAGCGGCAGCGTGAAATCAACT 120

Qy      242 TGATCGCTTACCGAGAAGCCCACAAT 267
      || || || || || || || || || ||
Db      121 TGTGGCCTATCATGGGGCATGCCAT 146

```

# RESULT 9

```

US-09-903-190-83
; Sequence 83, Application US/09903190
; Patent No. 6936692
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/903,190
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/247,155A
; PRIOR FILING DATE: 1999-02-09

```

```

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 83
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 612..644
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 829..834
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 850..861
US-09-903-190-83

```

```

Query Match          7.4%; Score 80.4; DB 3; Length 861;
Best Local Similarity 71.9%; Pred. No. 1.7e-11;
Matches 105; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

```

Qy      122 AACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTTT 181
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 AGCTCTGGTGGTTCTGGCTGCTCTGGACTGTCTCATCCTCTTTAGCTGCTGTTGCGCCT 60

Qy      182 GCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACC 241
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 TCCGCCACCGACGAGCTAAACTCAGGCTGCAACAACAGCAGCGGCAGCGTGAAATCAACT 120

Qy      242 TGATCGCTTACCGAGAAGCCCAACAAT 267
          | | | | | | | | | | | | | | | | |
Db      121 TGTGGCCTATCATGGGGCATGCCAT 146

```

# RESULT 10

US-09-148-545-55

```

; Sequence 55, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23

```

; EARLIER APPLICATION NUMBER: 60/047,502  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,633  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,583  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,617  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,618  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,503  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,592  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,581  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,584  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,500  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,587  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,492  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,598  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,613  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,582  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,596  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,612  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,632  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,601  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,580  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,568  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,314  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,569  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,311  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,671  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,674  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,669  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,312  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,313  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,672  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,315  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/048,974  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/056,886  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,877  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,889  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,893  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,630  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,878  
 ; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,631  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,845  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,892  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/047,595  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/057,761  
; EARLIER FILING DATE: 05-Sep-1997  
; EARLIER APPLICATION NUMBER: 60/047,599  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,588  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,585  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,586  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,590  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,594  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,589  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,593  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,614  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,578  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,576  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/047,501  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,670  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/056,632  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,664  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,876  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,881  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,909  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,875  
; EARLIER FILING DATE: 1997-08-22



; EARLIER APPLICATION NUMBER: 60/056,862  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,887  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,908  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/048,964  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/057,650  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/056,884  
 ; EARLIER FILING DATE: 1997-08-22  
 ; NUMBER OF SEQ ID NOS: 280  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 55  
 ; LENGTH: 1903

Query Match 4.0%; Score 43.2; DB 3; Length 1903;  
 Best Local Similarity 58.6%; Pred. No. 0.2;  
 Matches 75; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```
Qy      845 ACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACT 904
          ||||| || || | | ||||| |||| | | ||||| |||| |
Db      54 ACTTCTGCCACCGCGGGCCACCGCGGCCCGCTGATCCCGCAGAGGAAGGTCGCGGCCGTG 113

Qy      905 CCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCC 964
          | | | | | | ||||| ||||| ||||| ||||| |
Db      114 GAGCGATGACCCGCGCGGTCCGGCGGGCGCCCGGGGCTGCCACAGCCGCCCGCGCTTC 173

Qy      965 TGCTGCTG 972
          |||||
Db      174 TGCTGCTG 181
```

#### RESULT 11

US-09-621-011-55  
 ; Sequence 55, Application US/09621011  
 ; Patent No. 6878687  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 70 Human Secreted Proteins  
 ; FILE REFERENCE: P2001P1  
 ; CURRENT APPLICATION NUMBER: US/09/621,011  
 ; CURRENT FILING DATE: 2000-07-20  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 280  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 55  
 ; LENGTH: 1903  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-621-011-55

Query Match 4.0%; Score 43.2; DB 3; Length 1903;  
 Best Local Similarity 58.6%; Pred. No. 0.2;  
 Matches 75; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```
Qy      845 ACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACT 904
          ||||| || || | | ||||| |||| | | ||||| |||| |
Db      54 ACTTCTGCCACCGCGGGCCACCGCGGCCCGCTGATCCCGCAGAGGAAGGTCGCGGCCGTG 113

Qy      905 CCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCC 964
          | | | | | | ||||| ||||| ||||| ||||| |
Db      114 GAGCGATGACCCGCGCGGTCCGGCGGGCGCCCGGGGCTGCCACAGCCGCCCGCGCTTC 173

Qy      965 TGCTGCTG 972
          |||||
Db      174 TGCTGCTG 181
```

#### RESULT 12

US-09-252-991A-3601  
 ; Sequence 3601, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3601
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3601
```

```
Query Match          3.8%; Score 40.6; DB 3; Length 519;
Best Local Similarity 53.5%; Pred. No. 0.59;
Matches 85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
```

```
Qy      827 CTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGG 886
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      85  CCCTCGGCGAAGCGCTCGACAGCGGCGACATCCGCCGCCTGCTGCAACGCCAGGACGCCG 144

Qy      887 AGGAAGGCCTCTGTCACTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGC 946
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     -145 ACGACCCGATCCGCCAGTTGCACCGAGGAGCTGCGCGAGGGCCTCGCGCGCCTGGACCCGC 204

Qy      947 CACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAACG 985
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     205 GCGCGCTGCCCGGACGCGCCGTGCTGGACGCCTACGTG 243
```

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006
---------------------------------

# SCORE Search Results Details for Application 10729895 and Search Result 20070109\_162402\_us-10-729-895a-1.rnpbm.

[Score Home](#)    [Retrieve Application](#)    [SCORE System](#)    [SCORE](#)    [Comments /](#)  
[Page](#)    [List](#)    [Overview](#)    [FAQ](#)    [Suggestions](#)

This page gives you Search Results detail for the Application 10729895 and Search Result 20070109\_162402\_us-10-729-895a-1.rnpbm.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2007, 02:09:39 ; Search time 1616 Seconds  
(without alignments)  
8212.033 Million cell updates/sec

Title: US-10-729-895A-1  
Perfect score: 1080  
Sequence: 1 atgcctttccttttgggtct.....aaagcaaccagggtagggga 1080

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA\_Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

1	1080	100.0	1080	12	US-10-729-895-1	Sequence 1, Appli
2	1080	100.0	2976	7	US-10-297-880-7	Sequence 7, Appli
3	1080	100.0	4122	12	US-10-729-895-16	Sequence 16, Appl
4	1076.8	99.7	2703	7	US-10-094-749-823	Sequence 823, App
5	1054.4	97.6	1508	7	US-10-291-172-32	Sequence 32, Appl
6	1054.4	97.6	1508	8	US-10-221-278-32	Sequence 32, Appl
7	1054.4	97.6	1550	7	US-10-119-428-35	Sequence 35, Appl
8	1054.4	97.6	3095	3	US-09-814-353-20113	Sequence 20113, A
9	1053	97.5	1140	12	US-10-729-895-3	Sequence 3, Appli
10	1035.2	95.9	1715	7	US-10-291-172-408	Sequence 408, App
11	1035.2	95.9	1715	8	US-10-221-278-408	Sequence 408, App
12	783.2	72.5	69081	6	US-10-087-192-1192	Sequence 1192, Ap
c 13	505	46.8	505	3	US-09-864-761-8443	Sequence 8443, Ap
c 14	430.2	39.8	684973	3	US-09-263-959-1	Sequence 1, Appli
15	423	39.2	439	3	US-09-814-353-14563	Sequence 14563, A
16	402.4	37.3	566	4	US-09-925-065A-842189	Sequence 842189,
17	402.4	37.3	566	5	US-09-925-065A-842189	Sequence 842189,
c 18	352	32.6	352	3	US-09-864-761-25171	Sequence 25171, A
19	352	32.6	382	3	US-09-814-353-1832	Sequence 1832, Ap
20	352	32.6	382	3	US-09-814-353-8179	Sequence 8179, Ap
21	293	27.1	462	3	US-09-918-995-9876	Sequence 9876, Ap
22	231.6	21.4	399	10	US-10-926-683-82	Sequence 82, Appl
c 23	225.2	20.9	558	12	US-10-301-480-503459	Sequence 503459,
c 24	225.2	20.9	558	12	US-10-301-480-1116868	Sequence 1116868,
c 25	224	20.7	528	12	US-10-301-480-503460	Sequence 503460,
c 26	224	20.7	528	12	US-10-301-480-1116869	Sequence 1116869,
c 27	224	20.7	568	4	US-09-925-065A-442874	Sequence 442874,
c 28	224	20.7	568	5	US-09-925-065A-442874	Sequence 442874,
c 29	223.6	20.7	569	4	US-09-925-065A-442873	Sequence 442873,
c 30	223.6	20.7	569	5	US-09-925-065A-442873	Sequence 442873,
31	223	20.6	553	4	US-09-925-065A-409698	Sequence 409698,
32	223	20.6	553	5	US-09-925-065A-409698	Sequence 409698,
c 33	222.8	20.6	562	4	US-09-925-065A-125966	Sequence 125966,
c 34	222.8	20.6	562	5	US-09-925-065A-125966	Sequence 125966,
c 35	222.8	20.6	571	12	US-10-301-480-223349	Sequence 223349,
c 36	222.8	20.6	571	12	US-10-301-480-836758	Sequence 836758,
37	222.4	20.6	528	4	US-09-925-065A-393320	Sequence 393320,
38	222.4	20.6	528	4	US-09-925-065A-393321	Sequence 393321,
39	222.4	20.6	528	5	US-09-925-065A-393320	Sequence 393320,
40	222.4	20.6	528	5	US-09-925-065A-393321	Sequence 393321,
41	222.4	20.6	561	12	US-10-301-480-461406	Sequence 461406,
42	222.4	20.6	561	12	US-10-301-480-461407	Sequence 461407,
43	222.4	20.6	561	12	US-10-301-480-1074815	Sequence 1074815,
44	222.4	20.6	561	12	US-10-301-480-1074816	Sequence 1074816,
45	221.2	20.5	528	4	US-09-925-065A-393412	Sequence 393412,

#### ALIGNMENTS

RESULT 1  
 US-10-729-895-1  
 ; Sequence 1, Application US/10729895  
 ; Publication No. US20060063156A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: UNIVERSITY OF NEW MEXICO  
 ; TITLE OF INVENTION: OUTCOME PREDICTION AND RISK CLASSIFICATION IN CHILDHOOD  
 ; TITLE OF INVENTION: LEUKEMIA  
 ; FILE REFERENCE: N12-038US/310.00050101  
 ; CURRENT APPLICATION NUMBER: US/10/729,895  
 ; CURRENT FILING DATE: 2003-12-05  
 ; PRIOR APPLICATION NUMBER: 60/510,904  
 ; PRIOR FILING DATE: 2003-10-14  
 ; PRIOR APPLICATION NUMBER: 60/510,968  
 ; PRIOR FILING DATE: 2003-10-14  
 ; PRIOR APPLICATION NUMBER: 60/432,064  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: 60/432,077  
 ; PRIOR FILING DATE: 2002-12-06  
 ; PRIOR APPLICATION NUMBER: 60/432,078  
 ; PRIOR FILING DATE: 2002-12-06  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 1  
 ; LENGTH: 1080  
 ; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1026)  
US-10-729-895-1

Query Match 100.0%; Score 1080; DB 12; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	60
Db	1	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	60
Qy	61	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	120
Db	61	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	120
Qy	121	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	180
Db	121	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	180
Qy	181	TGCCACCACCGCCGAGCCAAGCACC GCCTTCAGGCCCAGCAGCGGCAACATGAAATCAAC	240
Db	181	TGCCACCACCGCCGAGCCAAGCACC GCCTTCAGGCCCAGCAGCGGCAACATGAAATCAAC	240
Qy	241	CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG	300
Db	241	CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTG	300
Qy	301	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA	360
Db	301	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA	360
Qy	361	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	420
Db	361	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	420
Qy	421	GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC	480
Db	421	GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC	480
Qy	481	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCT	540
Db	481	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCT	540
Qy	541	GACCTACCAAGTTGACCGAGCAGCCACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	600
Db	541	GACCTACCAAGTTGACCGAGCAGCCACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	600
Qy	601	GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	660
Db	601	GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	660
Qy	661	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	720
Db	661	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	720
Qy	721	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC	780
Db	721	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC	780
Qy	781	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	840
Db	781	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	840
Qy	841	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT	900
Db	841	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT	900
Qy	901	CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA	960
Db	901	CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA	960
Qy	961	TGCCTGCTGCTGAACACCATCAACAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC	1020

```

Db      961 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1020
QY      1021 CCCAGCTAGAGCAGGTCTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
          |||
Db      1021 CCCAGCTAGAGCAGGTCTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080

```

RESULT 2

```

US-10-297-880-7
; Sequence 7, Application US/10297880
; Publication No. US20030211513A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: HE, Ann; NGUYEN, Danniell B.
; APPLICANT: YAO, Monique G.; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; TANG, Y. Tom
; APPLICANT: XU, Yuming; HAFALIA, April J.A.
; APPLICANT: AZIMZAI, Yalda; CHAWLA, Narinder K.
; TITLE OF INVENTION: INTRACELLULAR SIGNALING PROTEINS
; FILE REFERENCE: PF-0782 USN
; CURRENT APPLICATION NUMBER: US/10/297,880
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US01/18595
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: US 60/210,582
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: US 60/212,443
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1478005CB1
US-10-297-880-7

```

```

Query Match      100.0%; Score 1080; DB 7; Length 2976;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60
          |||
Db      479 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 538
QY      61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120
          |||
Db      539 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 598
QY      121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180
          |||
Db      599 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 658
QY      181 TGCCACCACCGCCGAGCCAAGCACC GCCTTCAGGCCCAGCAGCGGCAACATGAAATCAAC 240
          |||
Db      659 TGCCACCACCGCCGAGCCAAGCACC GCCTTCAGGCCCAGCAGCGGCAACATGAAATCAAC 718
QY      241 CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTGG 300
          |||
Db      719 CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTGG 778
QY      301 CCAAACATTTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 360
          |||
Db      779 CCAAACATTTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 838
QY      361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 420
          |||
Db      839 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 898
QY      421 GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGACAGAGCAGCCCC 480
          |||
Db      899 GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGACAGAGCAGCCCC 958
QY      481 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540

```

```

Db      959  |||||TTGTCTGAGCCCAGCAGAAGCAGCACAAAGCCCCAAGCATCGCTGACCCTGATCCCTCT 1018
Qy      541  GACCTACCAAGTTGACCGAGCAGCCACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 600
Db      1019  |||||GACCTACCAAGTTGACCGAGCAGCCACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 1078
Qy      601  GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
Db      1079  |||||GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 1138
Qy      661  GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
Db      1139  |||||GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 1198
Qy      721  GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 780
Db      1199  |||||GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 1258
Qy      781  CACCATGACGATGACCTCAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
Db      1259  |||||CACCATGACGATGACCTCAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 1318
Qy      841  CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 900
Db      1319  |||||CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 1378
Qy      901  CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
Db      1379  |||||CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1438
Qy      961  TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1020
Db      1439  |||||TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1498
Qy      1021  CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
Db      1499  |||||CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1558

```

# RESULT 3

US-10-729-895-16

```

; Sequence 16, Application US/10729895
; Publication No. US20060063156A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF NEW MEXICO
; TITLE OF INVENTION: OUTCOME PREDICTION AND RISK CLASSIFICATION IN CHILDHOOD
; TITLE OF INVENTION: LEUKEMIA
; FILE REFERENCE: N12-038US/310.00050101
; CURRENT APPLICATION NUMBER: US/10/729,895
; CURRENT FILING DATE: 2003-12-05
; PRIOR APPLICATION NUMBER: 60/510,904
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/510,968
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/432,064
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 60/432,077
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 60/432,078
; PRIOR FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 16
; LENGTH: 4122
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-729-895-16

```

```

Query Match      100.0%; Score 1080; DB 12; Length 4122;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  ATGCCTTTCCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60
Db      145  ATGCCTTTCCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204

```

Qy	61	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	120
Db	205	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	264
Qy	121	GAAGTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	180
Db	265	GAAGTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	324
Qy	181	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	240
Db	325	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	384
Qy	241	CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTG	300
Db	385	CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTG	444
Qy	301	CCAAATATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA	360
Db	445	CCAAATATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA	504
Qy	361	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	420
Db	505	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	564
Qy	421	GGTGGCAGTCCCCCGGGCATCGATCCACACAGGGGATCCAGGGGGCACAGAGCAGCCCC	480
Db	565	GGTGGCAGTCCCCCGGGCATCGATCCACACAGGGGATCCAGGGGGCACAGAGCAGCCCC	624
Qy	481	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	540
Db	625	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	684
Qy	541	GACCTACCAGTTGACCGAGCAGCCACCAAAGCCCAGGGATGGAGCCAGTGGCTCTGTG	600
Db	685	GACCTACCAGTTGACCGAGCAGCCACCAAAGCCCAGGGATGGAGCCAGTGGCTCTGTG	744
Qy	601	GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	660
Db	745	GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	804
Qy	661	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	720
Db	805	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	864
Qy	721	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG	780
Db	865	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG	924
Qy	781	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC	840
Db	925	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCC	984
Qy	841	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT	900
Db	985	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT	1044
Qy	901	CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA	960
Db	1045	CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA	1104
Qy	961	TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC	1020
Db	1105	TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC	1164
Qy	1021	CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA	1080
Db	1165	CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA	1224

RESULT 4

US-10-094-749-823

; Sequence 823, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO



```
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 823
; LENGTH: 2703
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-823
```

```
Query Match          99.7%; Score 1076.8; DB 7; Length 2703;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 ATGCCTTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60
      |||
Db      145 ATGCCTTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204

Qy      61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120
      |||
Db      205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264

Qy      121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCATCCTGAGCTGCTGCTGTGTT 180
      |||
Db      265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCATCCTGAGCTGCTGCTGTGTT 324

Qy      181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240
      |||
Db      325 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384

Qy      241 CTGATCGCTTACCGAGAAGCCACAAATTACTCAGCGCTGCCATTTTATTTAGGTTTGTG 300
      |||
Db      385 CTGATCGCTTACCGAGAAGCCACAAATTACTCAGCGCTGCCATTTTATTTAGGTTTGTG 444

Qy      301 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 360
      |||
Db      445 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 504

Qy      361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 420
      |||
Db      505 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 564

Qy      421 GGTGGCAGTCCCCCGGGCATCGATCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCC 480
      |||
Db      565 GGTGGCAGTCCCCCGGGCATCGATCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCC 624

Qy      481 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540
      |||
Db      625 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 684

Qy      541 GACCTACCAGTTGACCGAGCAGCCACAAAGCCCGAGGATGGAGCCAGTGGCTCTGTG 600
      |||
Db      685 GACCTACCAGTTGACCGAGCAGCCACAAAGCCCGAGGATGGAGCCAGTGGCTCTGTG 744
```

```

Qy      601 GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
        |||
Db      745 GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 804

Qy      661 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
        |||
Db      805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864

Qy      721 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 780
        |||
Db      865 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924

Qy      781 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
        |||
Db      925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 984

Qy      841 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 900
        |||
Db      985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy      901 CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
        |||
Db      1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA 1104

Qy      961 TGCTGTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1020
        |||
Db      1105 TGCTGTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1164

Qy      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
        |||
Db      1165 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1224

```

# RESULT 5

```

US-10-291-172-32
; Sequence 32, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 32
; LENGTH: 1508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(1175)
US-10-291-172-32

```

```

Query Match      97.6%; Score 1054.4; DB 7; Length 1508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1055; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84
        |||
Db      174 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 233

Qy      85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
        |||

```

Db	234	TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG	293
Qy	145	TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC	204
Db	294	TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC	353
Qy	205	CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCCAC	264
Db	354	CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCCAC	413
Qy	265	AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAACTATTTACTACCTCCTTAT	324
Db	414	AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAACTATTTACTACCTCCTTAT	473
Qy	325	GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCAGCTACAG	384
Db	474	GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCAGCTACAG	533
Qy	385	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT	444
Db	534	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT	593
Qy	445	CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC	504
Db	594	CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC	653
Qy	505	ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC	564
Db	654	ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC	713
Qy	565	ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG	624
Db	714	ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG	773
Qy	625	GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT	684
Db	774	GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT	833
Qy	685	GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT	744
Db	834	GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT	893
Qy	745	GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG	804
Db	894	GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG	953
Qy	805	TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT	864
Db	954	TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT	1013
Qy	865	GTGCGGGCCCCCTGGTGTGATGAGGAGGAAGGCCTCTGTAGTCTCTGAGGAGCAGGCTCGA	924
Db	1014	GTGCGGGCCCCCTGGTGTGATGAGGAGGAAGGCCTCTGTAGTCTCTGAGGAGCAGGCTCGA	1073
Qy	925	GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC	984
Db	1074	GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC	1133
Qy	985	GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG	1044
Db	1134	GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG	1193
Qy	1045	CACCCAGCAACTTGGCAAAGCAACCAAGGGTAGGGGA	1080
Db	1194	CACCCAGCAACTTGGCAAAGCAACCAAGGGTAGGGGA	1229

# RESULT 6

US-10-221-278-32

; Sequence 32, Application US/10221278

; Publication No. US20040034208A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-045

; CURRENT APPLICATION NUMBER: US/10/221,278

```

; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 32
; LENGTH: 1508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (87)..(1175)
US-10-221-278-32

```

```

Query Match          97.6%; Score 1054.4; DB 8; Length 1508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1055; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84
      |||
Db      174 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 233
      |||

Qy      85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
      |||
Db      234 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 293
      |||

Qy      145 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGGCCACCACCGCCGAGCCAAGCAC 204
      |||
Db      294 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGGCCACCACCGCCGAGCCAAGCAC 353
      |||

Qy      205 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 264
      |||
Db      354 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 413
      |||

Qy      265 AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 324
      |||
Db      414 AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 473
      |||

Qy      325 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACAG 384
      |||
Db      474 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACAG 533
      |||

Qy      385 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 444
      |||
Db      534 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 593
      |||

Qy      445 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCCAGCAGAAGCAGC 504
      |||
Db      594 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCCAGCAGAAGCAGC 653
      |||

Qy      505 ACAAGACCCCCAAGCATCGCTGACCCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 564
      |||
Db      654 ACAAGACCCCCAAGCATCGCTGACCCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 713
      |||

Qy      565 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 624
      |||
Db      714 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 773
      |||

Qy      625 GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 684
      |||
Db      774 GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 833
      |||

Qy      685 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 744
      |||
Db      834 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 893
      |||

Qy      745 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 804

```

```

Db      894  |||||
GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 953

Qy      805  TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 864
|||||

Db      954  TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 1013

Qy      865  GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGA 924
|||||

Db      1014 GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGA 1073

Qy      925  GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 984
|||||

Db      1074 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1133

Qy      985  GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG 1044
|||||

Db      1134 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG 1193

Qy      1045 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
|||||

Db      1194 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1229

```

# RESULT 7

US-10-119-428-35

```

; Sequence 35, Application US/10119428
; Publication No. US20030165881A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Wehrman, Tom
; APPLICANT: Ren, Feiyan
; APPLICANT: Ma, Yunqing
; APPLICANT: Zhou, Ping
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030165881A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 789CIP2
; CURRENT APPLICATION NUMBER: US/10/119,428
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 35
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (129)..(1220)
US-10-119-428-35

```

```

Query Match          97.6%;  Score 1054.4;  DB 7;  Length 1550;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 1055;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

```

Qy      25  CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84
|||||

Db      216 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 275

Qy      85  TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAAGTCTGGTGGTTCTGGCTGGTG 144
|||||

Db      276 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAAGTCTGGTGGTTCTGGCTGGTG 335

Qy      145 TGGACCATCATCATCATCCTGAGTGCTGCTGCTGTTTGGCCACCACCGCCGAGCCAAGCAC 204
|||||

```

Db 336 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC 395

Qy 205 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 264  
|||||

Db 396 CGCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 455

Qy 265 AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 324  
|||||

Db 456 AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 515

Qy 325 GAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATACAGTGCCTTCCAGCTACAG 384  
|||||

Db 516 GAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATACAGTGCCTTCCAGCTACAG 575

Qy 385 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 444  
|||||

Db 576 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 635

Qy 445 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC 504  
|||||

Db 636 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC 695

Qy 505 ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 564  
|||||

Db 696 ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 755

Qy 565 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGAGCTGGACCCG 624  
|||||

Db 756 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGAGCTGGACCCG 815

Qy 625 GGGGCCTTCTTGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 684  
|||||

Db 816 GGGGCCTTCTTGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 875

Qy 685 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 744  
|||||

Db 876 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 935

Qy 745 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 804  
|||||

Db 936 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 995

Qy 805 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCAT 864  
|||||

Db 996 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCAT 1055

Qy 865 GTGCGGCCCCCTGGTGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCGA 924  
|||||

Db 1056 GTGCGGCCCCCTGGTGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCGA 1115

Qy 925 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 984  
|||||

Db 1116 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1175

Qy 985 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAG 1044  
|||||

Db 1176 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAG 1235

Qy 1045 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080  
|||||

Db 1236 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1271

RESULT 8

US-09-814-353-20113

; Sequence 20113, Application US/09814353

; Publication No. US20030165831A1

; GENERAL INFORMATION:

; APPLICANT: Lee, John

; APPLICANT: Thompson, Pamela

; APPLICANT: Lillie, James

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20113
; LENGTH: 3095
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 3094, 3095
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20113

```

```

Query Match          97.6%; Score 1054.4; DB 3; Length 3095;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1055; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      109 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 168

Qy      85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      169 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 228

Qy      145 TGGACCATCATCATCATCCTGAGCTGCTGCTGTTTGGCCACCACCGCCGAGCCAAGCAC 204
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      229 TGGACCATCATCATCATCCTGAGCTGCTGCTGTTTGGCCACCACCGCCGAGCCAAGCAC 288

Qy      205 CGCCTTCAGGCCCAGCAGCGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 264
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      289 CGCCTTCAGGCCCAGCAGCGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 348

Qy      265 AATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 324
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      349 AATTACTCAGCGCTGCCATTTTATTTTTCAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 408

Qy      325 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACAG 384
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      409 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACAG 468

Qy      385 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 444
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      469 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT 528

Qy      445 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCCAGCAGAAGCAGC 504
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      529 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCCAGCAGAAGCAGC 588

Qy      505 ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAAGTTGACCGAGCAGCC 564
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      589 ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAAGTTGACCGAGCAGCC 648

Qy      565 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 624
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      649 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 708

Qy      625 GGGGCCTTCTTGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 684
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      709 GGGGCCTTCTTGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 768

Qy      685 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGTTTACAGGT 744
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      769 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGTTTACAGGT 828

```

```

Qy      745 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 804
          |||
Db      829 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 888

Qy      805 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 864
          |||
Db      889 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 948

Qy      865 GTGCGGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGA 924
          |||
Db      949 GTGCGGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGA 1008

Qy      925 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 984
          |||
Db     1009 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1068

Qy      985 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG 1044
          |||
Db     1069 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG 1128

Qy     1045 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
          |||
Db     1129 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1164

```

RESULT 9

US-10-729-895-3

; Sequence 3, Application US/10729895

; Publication No. US20060063156A1

; GENERAL INFORMATION:

; APPLICANT: UNIVERSITY OF NEW MEXICO

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006
---------------------------------



# SCORE Search Results Details for Application 10729895 and Search Result 20070109\_162351\_us-10-729-895a-1.rng.

[Score Home](#)    [Retrieve Application](#)    [SCORE System](#)    [SCORE](#)    [Comments /](#)  
[Page](#)    [List](#)    [Overview](#)    [FAQ](#)    [Suggestions](#)

This page gives you Search Results detail for the Application 10729895 and Search Result 20070109\_162351\_us-10-729-895a-1.rng.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2007, 01:00:29 ; Search time 781 Seconds  
(without alignments)  
9641.522 Million cell updates/sec

Title: US-10-729-895A-1  
Perfect score: 1080  
Sequence: 1 atgcctttccttttgggtct.....aaagcaaccaggtagggga 1080

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_8:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*  
14: geneseqn2005s:\*  
15: geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1080	100.0	1080	12 ADP48574	Adp48574 Human OPA

2	1080	100.0	2976	6	AAI72319	Aai72319 ISIGP-2 c
3	1080	100.0	4122	12	ADP48589	Adp48589 Full leng
4	1076.8	99.7	2703	10	ADA53255	Ada53255 Human cod
5	1054.4	97.6	1508	5	AAS44951	Aas44951 cDNA enco
6	1054.4	97.6	3095	5	ADL61901	Adl61901 Human ova
7	1053	97.5	1140	12	ADP48576	Adp48576 Human OPA
8	1048	97.0	4150	4	AAD07816	Aad07816 Human sec
9	1043.4	96.6	4054	3	AAC59910	Aac59910 Human sec
10	1035.2	95.9	1715	5	AAS45139	Aas45139 cDNA enco
11	882	81.7	894	12	ADP28830	Adp28830 Human sec
12	789.8	73.1	4745	4	AAK65958	Aak65958 Human imm
13	788.2	73.0	4744	4	AAK65959	Aak65959 Human imm
14	783.2	72.5	69081	11	ACN44642	Acn44642 Human gen
c 15	664.6	61.5	832	4	AAD07857	Aad07857 Human sec
16	576.8	53.4	1006	5	ABA21206	Aba21206 Human ner
c 17	505	46.8	505	4	AAI17727	Aai17727 Probe #76
c 18	505	46.8	505	4	ABA62668	Aba62668 Human foe
c 19	505	46.8	505	4	AAI42669	Aai42669 Probe #11
c 20	505	46.8	505	4	ABA29977	Aba29977 Probe #84
c 21	505	46.8	505	4	AAK36874	Aak36874 Human bon
c 22	505	46.8	505	4	AAK11046	Aak11046 Human bra
c 23	505	46.8	505	4	ABS36548	Abs36548 Human liv
c 24	505	46.8	505	6	ABS10887	Abs10887 Human gen
c 25	430.2	39.8	84973	12	ADH69807_6	Continuation (7 of
26	423	39.2	439	5	ADL40673	Adl40673 Human ova
27	384.6	35.6	577	8	ABZ36387	Abz36387 Human sec
28	365.4	33.8	550	4	AAH98928	Aah98928 Human EST
29	352.2	32.6	605	4	AAD07856	Aad07856 Human sec
c 30	352	32.6	352	4	AAI26927	Aai26927 Probe #16
c 31	352	32.6	352	4	ABA75183	Aba75183 Human foe
c 32	352	32.6	352	4	AAI55731	Aai55731 Probe #24
c 33	352	32.6	352	4	ABA39851	Aba39851 Probe #18
c 34	352	32.6	352	4	AAK49818	Aak49818 Human bon
c 35	352	32.6	352	4	AAK23717	Aak23717 Human bra
c 36	352	32.6	352	4	ABS49458	Abs49458 Human liv
c 37	352	32.6	352	6	ABS23320	Abs23320 Human gen
38	352	32.6	382	5	ADI75437	Adi75437 Human ova
39	352	32.6	382	5	ADI69090	Adi69090 Human ova
40	330.4	30.6	435	13	ADQ53832	Adq53832 Novel can
41	321.8	29.8	603	13	ADU02259	Adu02259 Novel hum
42	304.8	28.2	489	5	ABA14114	Aba14114 Human ner
43	293	27.1	462	9	ACH22664	Ach22664 Human adu
44	231.6	21.4	399	3	AAZ42323	Aaz42323 Human 5'
45	231.6	21.4	399	3	AAC00276	Aac00276 Human sec

#### ALIGNMENTS

##### RESULT 1

ADP48574

ID ADP48574 standard; DNA; 1080 BP.

XX

AC ADP48574;

XX

DT 09-SEP-2004 (first entry)

XX

DE Human OPAL1/G0 DNA splice form incorporating exon 1 SeqID 1.

XX

KW human; gene; ds; OPAL1; G0; childhood leukaemia; G protein beta 2; G1;

KW interleukin-10 receptor alpha; G2; IL-10; FYN-binding protein; PBK1;

KW cytostatic; acute lymphoblastic leukaemia; ALL; acute myeloid leukaemia;

KW AML.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1029.

FT /\*tag= a

FT /product= "OPAL1/G0 protein (incorporating exon 1)"

XX

PN WO2004053074-A2.

XX

PD 24-JUN-2004.

XX

PF 05-DEC-2003; 2003WO-US038738.

XX  
PR 06-DEC-2002; 2002US-0432064P.  
PR 06-DEC-2002; 2002US-0432077P.  
PR 06-DEC-2002; 2002US-0432078P.  
PR 14-OCT-2003; 2003US-0510904P.  
PR 14-OCT-2003; 2003US-0510968P.  
PR 05-DEC-2003; 2003US-00729895.  
PR 05-DEC-2003; 2003US-0527610P.  
XX  
PA (SCTE-) SCI & TECHNOLOGY CORP @UNM.  
PA (SAND-) SANDIA CORP.  
XX  
PI Willman CL, Helman P, Veroff R, Mosquera-Caro M, Davidson GS;  
PI Martin SB, Atlas SR, Andries E, Kang H, Shuster JJ, Haaland DM;  
PI Potter JW, Wang X, Harvey R;  
XX  
DR WPI; 2004-468846/44.  
DR P-PSDB; ADP48575.  
XX  
PT New isolated Outcome Predictor in Acute Leukemia 1 (OPAL1)  
PT polynucleotides and polypeptides for preventing, treating or diagnosing  
PT leukemia, or for screening or evaluating compounds that may be used for  
PT treating leukemia.  
XX  
PS Claim 1; SEQ ID NO 1; 343pp; English.  
XX  
CC This invention relates to a novel isolated nucleic acid identified as  
CC OPAL1 (also known as G0), and the encoded protein thereof. Specifically,  
CC it refers to a method for predicting the outcome and risk classification  
CC of childhood leukaemia. The present invention describes a gene profiling  
CC method that can be used to determine the expression level for OPAL1, a  
CC gene that is strongly predictive regarding the outcome of both acute  
CC lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML).  
CC Furthermore, it provides additional genes namely G protein beta 2  
CC (referred to as G1), interleukin-10 (IL-10) receptor alpha (referred to  
CC as G2), the FYN-binding protein and PBK1 that can also be used, alone or  
CC in combination, in predictive assays for the classification and potential  
CC therapy for a leukaemia patient. Accordingly, these compositions exhibit  
CC cytostatic activities. This polynucleotide sequence is the human OPAL1/G0  
CC splice from incorporating exon 1 of the invention.  
XX  
SQ Sequence 1080 BP; 249 A; 341 C; 292 G; 198 T; 0 U; 0 Other;

Query Match 100.0%; Score 1080; DB 12; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 2.9e-278;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCCTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60  
|||||  
Db 1 ATGCCCTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60

Qy 61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120  
|||||  
Db 61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120

Qy 121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180  
|||||  
Db 121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180

Qy 181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240  
|||||  
Db 181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240

Qy 241 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 300  
|||||  
Db 241 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 300

Qy 301 CCAAATACTTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 360  
|||||  
Db 301 CCAAATACTTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 360

Qy 361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCA 420  
|||||  
Db 361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCA 420

Qy 421 GGTGGCAGTCCCCCGGCATCGATCCACCAGGGGATCCAGGGGGCACAGAGCAGCCCC 480

```

|||||
Db      421  GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC 480
Qy      481  TTGTCTGAGCCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540
|||||
Db      481  TTGTCTGAGCCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540
Qy      541  GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 600
|||||
Db      541  GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 600
Qy      601  GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
|||||
Db      601  GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
Qy      661  GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
|||||
Db      661  GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
Qy      721  GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG 780
|||||
Db      721  GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG 780
Qy      781  CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
|||||
Db      781  CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
Qy      841  CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT 900
|||||
Db      841  CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT 900
Qy      901  CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
|||||
Db      901  CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
Qy      961  TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1020
|||||
Db      961  TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1020
Qy      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
|||||
Db      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080

```

# RESULT 2

AAI72319

ID AAI72319 standard; cDNA; 2976 BP.

XX

AC AAI72319;

XX

DT 15-APR-2002 (first entry)

XX

DE ISIGP-2 cDNA.

XX

KW Human; intracellular signalling protein; ISIGP; gene; cell proliferation;  
KW autoimmune; inflammation; gastrointestinal disorder;  
KW reproductive disorder; developmental disorder; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 479..1507

FT /\*tag= a

FT /product= "ISIGP-2"

XX

PN WO200194391-A2.

XX

PD 13-DEC-2001.

XX

PF 07-JUN-2001; 2001WO-US018595.

XX

PR 08-JUN-2000; 2000US-0210582P.

PR 16-JUN-2000; 2000US-0212443P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Yue H, He A, Nguyen DB, Yao MG, Bandman O, Burford N, Tang YT;  
 PI Xu Y, Hafalia A, Azimzai Y, Walia NK;  
 XX  
 DR WPI; 2002-154564/20.  
 DR P-PSDB; AAB47872.  
 XX  
 PT New human intracellular signaling protein and polynucleotides useful for  
 PT diagnosing, treating or preventing cell proliferative,  
 PT autoimmune/inflammatory, gastrointestinal, reproductive and developmental  
 PT disorders.  
 XX  
 PS Claim 5; Page 102-03; 106pp; English.  
 XX  
 CC The sequences given in AAI72318-22 encode novel human intracellular  
 CC signalling proteins (ISIGP). The polynucleotides and ISIGP proteins may  
 CC be used for the diagnosis, treatment or prevention of cell proliferative,  
 CC autoimmune/inflammatory, gastrointestinal, reproductive and developmental  
 CC disorders. The protein encoded by this sequence has homology to human WW  
 CC domain binding protein-1  
 XX  
 SQ Sequence 2976 BP; 633 A; 817 C; 823 G; 703 T; 0 U; 0 Other;

Query Match 100.0%; Score 1080; DB 6; Length 2976;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-278;  
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	60
Db	479	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	538
Qy	61	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	120
Db	539	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	598
Qy	121	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	180
Db	599	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	658
Qy	181	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	240
Db	659	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	718
Qy	241	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAAGTTTTTG	300
Db	719	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAAGTTTTTG	778
Qy	301	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA	360
Db	779	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA	838
Qy	361	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	420
Db	839	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	898
Qy	421	GGTGGCAGTCCCCGGGCATCGATCCCACCAGGGGATCCAGGGGGCACAGAGCAGCCCC	480
Db	899	GGTGGCAGTCCCCGGGCATCGATCCCACCAGGGGATCCAGGGGGCACAGAGCAGCCCC	958
Qy	481	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	540
Db	959	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	1018
Qy	541	GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCAAGGATGGAGCCAGTGGCTCTGTG	600
Db	1019	GACCTACCAAGTTGACCGAGCAGCCACCAAGCCCAAGGATGGAGCCAGTGGCTCTGTG	1078
Qy	601	GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	660
Db	1079	GCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	1138
Qy	661	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	720
Db	1139	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	1198
Qy	721	GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGAACCGGGGC	780

```

Db      1199 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 1258
QY      781 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1259 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 1318
QY      841 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1319 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1378
QY      901 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1379 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1438
QY      961 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1439 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1498
QY      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1499 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1558

```

RESULT 3

ADP48589

ID ADP48589 standard; cDNA; 4122 BP.

XX

AC ADP48589;

XX

DT 09-SEP-2004 (first entry)

XX

DE Full length human OPAL1 cDNA SeqID 16.

XX

KW human; ss; OPAL1; G0; childhood leukaemia; G protein beta 2; G1;

KW interleukin-10 receptor alpha; G2; IL-10; FYN-binding protein; PBK1;

KW cytostatic; acute lymphoblastic leukaemia; ALL; acute myeloid leukaemia;

KW AML.

XX

OS Homo sapiens.

XX

PN WO2004053074-A2.

XX

PD 24-JUN-2004.

XX

PF 05-DEC-2003; 2003WO-US038738.

XX

PR 06-DEC-2002; 2002US-0432064P.

PR 06-DEC-2002; 2002US-0432077P.

PR 06-DEC-2002; 2002US-0432078P.

PR 14-OCT-2003; 2003US-0510904P.

PR 14-OCT-2003; 2003US-0510968P.

PR 05-DEC-2003; 2003US-00729895.

PR 05-DEC-2003; 2003US-0527610P.

XX

PA (SCTE-) SCI & TECHNOLOGY CORP @UNM.

PA (SAND-) SANDIA CORP.

XX

PI Willman CL, Helman P, Veroff R, Mosquera-Caro M, Davidson GS;

PI Martin SB, Atlas SR, Andries E, Kang H, Shuster JJ, Haaland DM;

PI Potter JW, Wang X, Harvey R;

XX

DR WPI; 2004-468846/44.

XX

PT New isolated Outcome Predictor in Acute Leukemia 1 (OPAL1)

PT polynucleotides and polypeptides for preventing, treating or diagnosing

PT leukemia, or for screening or evaluating compounds that may be used for

PT treating leukemia.

XX

PS Disclosure; SEQ ID NO 16; 343pp; English.

XX

CC This invention relates to a novel isolated nucleic acid identified as  
 CC OPAL1 (also known as G0), and the encoded protein thereof. Specifically,  
 CC it refers to a method for predicting the outcome and risk classification  
 CC of childhood leukaemia. The present invention describes a gene profiling  
 CC method that can be used to determine the expression level for OPAL1, a  
 CC gene that is strongly predictive regarding the outcome of both acute

CC lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML).  
 CC Furthermore, it provides additional genes namely G protein beta 2  
 CC (referred to as G1), interleukin-10 (IL-10) receptor alpha (referred to  
 CC as G2), the FYN-binding protein and PBK1 that can also be used, alone or  
 CC in combination, in predictive assays for the classification and potential  
 CC therapy for a leukaemia patient. Accordingly, these compositions exhibit  
 CC cytostatic activities. This polynucleotide sequence is the full length  
 CC human OPAL1 cDNA of the invention.

XX

SQ Sequence 4122 BP; 964 A; 1066 C; 1075 G; 1017 T; 0 U; 0 Other;

Query Match 100.0%; Score 1080; DB 12; Length 4122;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-278;  
 Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	60
Db	145	ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA	204
QY	61	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	120
Db	205	AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT	264
QY	121	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	180
Db	265	GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT	324
QY	181	TGCCACCACCGCCGAGCCAAGCACCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	240
Db	325	TGCCACCACCGCCGAGCCAAGCACCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	384
QY	241	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAAGTTTTGG	300
Db	385	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAAGTTTTGG	444
QY	301	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA	360
Db	445	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA	504
QY	361	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	420
Db	505	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	564
QY	421	GGTGGCAGTCCCCCGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC	480
Db	565	GGTGGCAGTCCCCCGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC	624
QY	481	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	540
Db	625	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	684
QY	541	GACCTACAGTTGACCGAGCAGCCACCAAGCCCAAGGGATGGAGCCAGTGGCTCTGTG	600
Db	685	GACCTACAGTTGACCGAGCAGCCACCAAGCCCAAGGGATGGAGCCAGTGGCTCTGTG	744
QY	601	GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	660
Db	745	GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	804
QY	661	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	720
Db	805	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	864
QY	721	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG	780
Db	865	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGG	924
QY	781	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	840
Db	925	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	984
QY	841	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT	900
Db	985	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTTGGTGATGAGGAGGAAGGCCTCTGT	1044
QY	901	CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA	960

```

|||||
Db      1045 CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1104
Qy      961 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1020
|||||
Db      1105 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1164
Qy      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
|||||
Db      1165 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1224

```

#### RESULT 4

ADA53255

ID ADA53255 standard; cDNA; 2703 BP.

XX

AC ADA53255;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human coding sequence, SEQ ID 823.

XX

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

KW Gene Therapy; human; secretory protein; membrane proteins; cancer;

KW inflammatory disease; osteoporosis; neurological disease; gene; ss.

XX

OS Homo sapiens.

XX

PN EP1293569-A2.

XX

PD 19-MAR-2003.

XX

PF 21-MAR-2002; 2002EP-00006586.

XX

PR 14-SEP-2001; 2001JP-00328381.

PR 24-JAN-2002; 2002US-0350435P.

XX

PA (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX

DR WPI; 2003-395539/38.

DR P-PSDB; ADA54894.

XX

PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.

XX

PS Claim 1; SEQ ID NO 823; 205pp; English.

XX

CC The present invention relates to novel human secretory or membrane

CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-

CC ADA54071). The coding sequences are useful in the gene therapy of

CC diseases caused by abnormalities of the proteins, e.g. cancer,

CC inflammatory diseases, osteoporosis or neurological disease.

XX

SQ Sequence 2703 BP; 599 A; 742 C; 713 G; 649 T; 0 U; 0 Other;

Query Match 99.7%; Score 1076.8; DB 10; Length 2703;

Best Local Similarity 99.8%; Pred. No. 3.1e-277;

Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      1 ATGCCTTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60
|||||
Db      145 ATGCCTTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204
Qy      61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120
|||||
Db      205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264
Qy      121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180
|||||
Db      265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 324

```



Qy	181	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	240
Db	325	TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC	384
Qy	241	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG	300
Db	385	CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG	444
Qy	301	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA	360
Db	445	CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCA	504
Qy	361	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	420
Db	505	CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA	564
Qy	421	GGTGGCAGTCCCCCGGCATCGATCCACACAGGGGATCCAGGGGGCACAGAGCAGCCCC	480
Db	565	GGTGGCAGTCCCCCGGCATCGATCCACACAGGGGATCCAGGGGGCACAGAGCAGCCCC	624
Qy	481	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	540
Db	625	TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT	684
Qy	541	GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	600
Db	685	GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG	744
Qy	601	GCTGGCCTGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	660
Db	745	GCTGGCCTGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG	804
Qy	661	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	720
Db	805	GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT	864
Qy	721	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGC	780
Db	865	GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGC	924
Qy	781	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	840
Db	925	CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC	984
Qy	841	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT	900
Db	985	CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCTGGTGATGAGGAGGAAGGCCTCTGT	1044
Qy	901	CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA	960
Db	1045	CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA	1104
Qy	961	TGCCTGCTGCTGAACACCATCAACAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC	1020
Db	1105	TGCCTGCTGCTGAACACCATCAACAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC	1164
Qy	1021	CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA	1080
Db	1165	CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA	1224

RESULT 5

AAS44951

ID AAS44951 standard; cDNA; 1508 BP.

XX

AC AAS44951;

XX

DT 18-DEC-2001 (first entry)

XX

DE cDNA encoding novel human secretory protein, Seq ID No 32.

XX

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;

KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;

KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;

KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;

KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
 KW fertility; analgesic; pain; antigen; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200166689-A2.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 05-MAR-2001; 2001WO-US004942.  
 XX  
 PR 07-MAR-2000; 2000US-00519705.  
 PR 19-MAY-2000; 2000US-00574454.  
 PR 17-JUN-2000; 2000US-00596193.  
 PR 14-JUL-2000; 2000US-00616847.  
 PR 19-SEP-2000; 2000US-00665363.  
 PR 20-OCT-2000; 2000US-00693267.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;  
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;  
 XX  
 DR WPI; 2001-589934/66.  
 DR P-PSDB; AAU28051.  
 XX  
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues, for diagnosis and treatment of  
 PT cancer, neurological, inflammatory, and autoimmune disorders.  
 XX  
 PS Claim 1; SEQ ID NO 32; 107pp; English.  
 XX  
 CC The invention relates to novel isolated human secreted polypeptides (I)  
 CC and polynucleotides (II). (I) and (II) are useful for treating  
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
 CC involved in increasing haematopoiesis, stem cell survival, bone growth  
 CC and remodeling. (I), (II) and modulators of (II) are useful for  
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
 CC creating transgenic animals useful for studying the in vivo activities of  
 CC the polypeptide as well as for studying modulators of the polypeptides.  
 CC (I) induces the proliferation of neural cells and regeneration of nerve  
 CC and brain tissue and is useful for the treatment of central and  
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,  
 CC or periodontal disease. Furthermore, (I) is also useful for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues, various immune deficiencies and  
 CC disorders including severe combined immunodeficiency (SCID), bacterial or  
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,  
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
 CC reactions and conditions, such as asthma or other respiratory problems.  
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,  
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
 CC analgesic effects or other pain reducing effects, immunoglobulin like  
 CC activity and can act as an antigen in a vaccine composition to raise an  
 CC immune response. AAS44920-AAS45295 represent novel human secreted protein  
 CC coding sequences of the invention  
 XX  
 SQ Sequence 1508 BP; 345 A; 447 C; 423 G; 293 T; 0 U; 0 Other;

Query Match 97.6%; Score 1054.4; DB 5; Length 1508;  
 Best Local Similarity 99.9%; Pred. No. 2.4e-271;  
 Matches 1055; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84

```

      |||
Db      174 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 233
QY      85 TGCTGTGGACAGTCTCAGTGTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
      |||
Db      234 TGCTGTGGACAGTCTCAGTGTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 293
QY      145 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGGCCACCACCGCCGAGCCAAGCAC 204
      |||
Db      294 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTTGGCCACCACCGCCGAGCCAAGCAC 353
QY      205 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 264
      |||
Db      354 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 413
QY      265 AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 324
      |||
Db      414 AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAAACTATTTACTACCTCCTTAT 473
QY      325 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCAGCTACAG 384
      |||
Db      474 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCAGCTACAG 533
QY      385 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT 444
      |||
Db      534 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT 593
QY      445 CCCACCAGGGGATCCCAGGGGACACAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC 504
      |||
Db      594 CCCACCAGGGGATCCCAGGGGACACAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGC 653
QY      505 ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 564
      |||
Db      654 ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 713
QY      565 ACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 624
      |||
Db      714 ACCAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 773
QY      625 GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 684
      |||
Db      774 GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 833
QY      685 GAACACGGCGACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 744
      |||
Db      834 GAACACGGCGACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT 893
QY      745 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 804
      |||
Db      894 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 953
QY      805 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT 864
      |||
Db      954 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT 1013
QY      865 GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCGA 924
      |||
Db      1014 GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCGA 1073
QY      925 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 984
      |||
Db      1074 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1133
QY      985 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG 1044
      |||
Db      1134 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAG 1193
QY      1045 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
      |||
Db      1194 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1229

```

RESULT 6

ADL61901

ID ADL61901 standard; DNA; 3095 BP.

XX

AC ADL61901;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human ovarian cancer DNA marker #20113.  
 XX  
 KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200170979-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US009126.  
 XX  
 PR 21-MAR-2000; 2000US-0191031P.  
 PR 25-MAY-2000; 2000US-0207124P.  
 PR 15-JUN-2000; 2000US-0211940P.  
 PR 07-JUL-2000; 2000US-0216820P.  
 PR 25-JUL-2000; 2000US-0220661P.  
 PR 21-DEC-2000; 2000US-0257672P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Lee J, Lillie J;  
 XX  
 DR WPI; 2001-611502/70.  
 XX  
 PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
 PT cancer cells as compared to their normal non-cancerous ovarian cells are  
 PT used to characterize stage, grade, histological type of ovarian cancer.  
 XX  
 PS Disclosure; SEQ ID NO 20113; 106pp; English.  
 XX  
 CC The invention relates to nucleic acid markers which are overexpressed in  
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-  
 CC cancerous) ovarian cells. The invention also relates to polypeptides  
 CC encoded by the markers, antibodies that selectively bind to the  
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
 CC of developing ovarian cancer involving inhibiting expression of a gene  
 CC corresponding to a marker of the invention and a method of treating a  
 CC patient afflicted with ovarian cancer comprising providing to cells of  
 CC the patient an antisense oligonucleotide complementary to a marker of the  
 CC invention. The markers are useful for assessing if a patient is afflicted  
 CC with ovarian cancer, which involves comparing the level of expression of  
 CC a marker in a patient sample and a normal level of expression of the  
 CC marker in a control non-ovarian cancer sample. A difference between the  
 CC expression levels indicates ovarian cancer. The level of expression of a  
 CC marker corresponds to a secreted protein or to a transcribed  
 CC polynucleotide or its portion. The level of expression of the marker is  
 CC assessed by detecting the presence in the sample, a protein or protein  
 CC fragment corresponding to the marker. The presence of protein or protein  
 CC fragment is detected using an antibody that specifically binds with the  
 CC protein or protein fragment. Alternatively, the level of expression of  
 CC the marker is assessed by detecting the presence of a transcribed  
 CC polynucleotide which anneals with the marker or anneals with a portion of  
 CC the polynucleotide comprising the marker, under stringent conditions. The  
 CC marker is also used for monitoring the progression of ovarian cancer in a  
 CC patient which involves detecting expression of the marker in a patient  
 CC sample at a first point in time, repeating the method at a subsequent  
 CC time and comparing the level of expression. The method is carried out  
 CC using an ovarian tissue sample. A composition comprising a marker,  
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
 CC This sequence represents a human ovarian cancer DNA marker of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 3095 BP; 702 A; 832 C; 845 G; 711 T; 0 U; 5 Other;

Query Match 97.6%; Score 1054.4; DB 5; Length 3095;  
 Best Local Similarity 99.9%; Pred. No. 3.2e-271;  
 Matches 1055; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84

```

|||||
Db      109 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 168
Qy      85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144
|||||
Db      169 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 228
Qy      145 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC 204
|||||
Db      229 TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC 288
Qy      205 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 264
|||||
Db      289 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 348
Qy      265 AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACTATTTACTACCTCCTTAT 324
|||||
Db      349 AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACTATTTACTACCTCCTTAT 408
Qy      325 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCAGCTACAG 384
|||||
Db      409 GAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCAGCTACAG 468
Qy      385 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT 444
|||||
Db      469 CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGCATCGAT 528
Qy      445 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCCCTGTCTGAGCCAGCAGAAGCAGC 504
|||||
Db      529 CCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCCCTGTCTGAGCCAGCAGAAGCAGC 588
Qy      505 ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 564
|||||
Db      589 ACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC 648
Qy      565 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 624
|||||
Db      649 ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCG 708
Qy      625 GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 684
|||||
Db      709 GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT 768
Qy      685 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGTTTACAGGT 744
|||||
Db      769 GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGTTTACAGGT 828
Qy      745 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 804
|||||
Db      829 GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG 888
Qy      805 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 864
|||||
Db      889 TTCAACACACTCATCGATGATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCAT 948
Qy      865 GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCGA 924
|||||
Db      949 GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCGA 1008
Qy      925 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 984
|||||
Db      1009 GAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC 1068
Qy      985 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAG 1044
|||||
Db      1069 GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAG 1128
Qy      1045 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
|||||
Db      1129 CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1164

```

RESULT 7  
ADP48576

ID ADP48576 standard; DNA; 1140 BP.  
XX

AC ADP48576;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE Human OPAL1/G0 DNA splice form incorporating exon 1a SeqID 3.  
 XX  
 KW human; gene; ds; OPAL1; G0; childhood leukaemia; G protein beta 2; G1;  
 KW interleukin-10 receptor alpha; G2; IL-10; FYN-binding protein; PBK1;  
 KW cytostatic; acute lymphoblastic leukaemia; ALL; acute myeloid leukaemia;  
 KW AML.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1092  
 FT /\*tag= a  
 FT /product= "OPAL1/G0 protein (incorporating exon 1a)"  
 XX  
 PN WO2004053074-A2.  
 XX  
 PD 24-JUN-2004.  
 XX  
 PF 05-DEC-2003; 2003WO-US038738.  
 XX  
 PR 06-DEC-2002; 2002US-0432064P.  
 PR 06-DEC-2002; 2002US-0432077P.  
 PR 06-DEC-2002; 2002US-0432078P.  
 PR 14-OCT-2003; 2003US-0510904P.  
 PR 14-OCT-2003; 2003US-0510968P.  
 PR 05-DEC-2003; 2003US-00729895.  
 PR 05-DEC-2003; 2003US-0527610P.  
 XX

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006
---------------------------------

<a href="#">Score Home</a>	<a href="#">Retrieve Application</a>	<a href="#">SCORE System</a>	<a href="#">SCORE</a>	<a href="#">Comments /</a>
<a href="#">Page</a>	<a href="#">List</a>	<a href="#">Overview</a>	<a href="#">FAQ</a>	<a href="#">Suggestions</a>

[start](#) | [next page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2007 Bioceleration Ltd.

## SUMMARIES

http://es/ScoreAccessWeb/GetItem.action?AppId=10729895&seqId=1038736&ItemName... 1/22/2007

2	1076.8	99.7	2703	2	AX714139	AX714139 Sequence
3	1076.8	99.7	2703	5	AK056285	AK056285 Homo sapi
4	1054.4	97.6	3095	2	CQ413042	CQ413042 Sequence
5	1043.4	96.6	4054	2	BD275013	BD275013 50 Human
6	1002	92.8	1002	2	CQ720290	CQ720290 Sequence
7	789.8	73.1	131753	5	AL358790	AL358790 Human DNA
8	789.8	73.1	173758	5	AC009144	AC009144 Homo sapi
9	783.2	72.5	62035	5	AC078937	AC078937 Homo sapi
10	777.6	72.0	170261	5	AC147974	AC147974 Pan trogl
11	776.2	71.9	228263	12	AC079930	AC079930 Homo sapi
12	770	71.3	2241	6	BC026369	BC026369 Mus muscu
13	766.8	71.0	3921	6	BC058949	BC058949 Mus muscu
14	703.4	65.1	2217	5	AK000374	AK000374 Homo sapi
15	587.8	54.4	150528	12	AC149242	AC149242 Macaca mu
16	587.8	54.4	166044	12	AC150715	AC150715 Macaca mu
c 17	587.8	54.4	176436	12	AC148535	AC148535 Macaca mu
18	576.6	53.4	275121	12	AC099420	AC099420 Rattus no
c 19	519.2	48.1	74874	12	AC168985	AC168985 Bos tauru
20	518.8	48.0	196003	12	AC079419	AC079419 Mus muscu
c 21	518.8	48.0	206019	6	AC161865	AC161865 Mus muscu
22	518.8	48.0	216311	6	AC156982	AC156982 Mus muscu
c 23	505	46.8	505	2	CQ071860	CQ071860 Sequence
c 24	505	46.8	505	2	CQ102496	CQ102496 Sequence
c 25	505	46.8	505	2	CQ141409	CQ141409 Sequence
c 26	505	46.8	505	2	CQ177047	CQ177047 Sequence
c 27	505	46.8	505	2	CQ224699	CQ224699 Sequence
c 28	505	46.8	505	2	CQ262712	CQ262712 Sequence
c 29	505	46.8	505	2	CQ299773	CQ299773 Sequence
c 30	505	46.8	505	2	CQ336943	CQ336943 Sequence
c 31	453.8	42.0	239579	12	AC095597	AC095597 Rattus no
c 32	453.8	42.0	247659	12	AC095548	AC095548 Rattus no
33	449	41.6	247169	12	AC123007	AC123007 Rattus no
34	430.2	39.8	28657	12	AC153482	AC153482 Homo sapi
c 35	430.2	39.8	35944	5	HSTCRB1H	AF009664 Homo sapi
36	430.2	39.8	67826	5	AC099395	AC099395 Homo sapi
c 37	430.2	39.8	161425	5	AC104597	AC104597 Homo sapi
c 38	430.2	39.8	232650	5	U66061	U66061 Human germl
39	423	39.2	439	2	CQ407492	CQ407492 Sequence
c 40	352	32.6	352	2	CQ081060	CQ081060 Sequence
c 41	352	32.6	352	2	CQ115558	CQ115558 Sequence
c 42	352	32.6	352	2	CQ154353	CQ154353 Sequence
c 43	352	32.6	352	2	CQ186921	CQ186921 Sequence
c 44	352	32.6	352	2	CQ237609	CQ237609 Sequence
c 45	352	32.6	352	2	CQ275227	CQ275227 Sequence

#### ALIGNMENTS

RESULT 1  
AX337966  
LOCUS AX337966 2976 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 7 from Patent WO0194391.  
ACCESSION AX337966  
VERSION AX337966.1 GI:18128677  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Yue,H., He,A., Nguyen,D.B., Yao,M.G., Bandman,O., Burford,N.,  
Tang,Y.T., Xu,Y., Hafalia,A., Azimzai,Y. and Walia,N.K.  
TITLE Intracellular signaling proteins  
JOURNAL Patent: WO 0194391-A 7 13-DEC-2001;  
Incyte Genomics, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1. .2976  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 1478005CB1"  
ORIGIN  
Query Match 100.0%; Score 1080; DB 2; Length 2976;



Best Local Similarity 100.0%; Pred. No. 1.4e-310;  
Matches 1080; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60
      |||
Db      479 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 538

Qy      61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120
      |||
Db      539 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 598

Qy      121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180
      |||
Db      599 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 658

Qy      181 TGCCACCACC GCCGAGCCAAGCACC GCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240
      |||
Db      659 TGCCACCACC GCCGAGCCAAGCACC GCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 718

Qy      241 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTAGGTTTTTG 300
      |||
Db      719 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTAGGTTTTTG 778

Qy      301 CCAAATATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA 360
      |||
Db      779 CCAAATATTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCCA 838

Qy      361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 420
      |||
Db      839 CCATACAGTGCCTTCCAGCTACAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 898

Qy      421 GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC 480
      |||
Db      899 GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC 958

Qy      481 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540
      |||
Db      959 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 1018

Qy      541 GACCTACCAAGTTGACCGAGCAGCCACAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 600
      |||
Db      1019 GACCTACCAAGTTGACCGAGCAGCCACAAAGCCCCAGGGATGGAGCCAGTGGCTCTGTG 1078

Qy      601 GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
      |||
Db      1079 GCTGGCCTGGGGGAGCTGGACCCGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 1138

Qy      661 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
      |||
Db      1139 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 1198

Qy      721 GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGCAACCGGGGC 780
      |||
Db      1199 GGGAGACATCGCCGCTTACAGGTGACTCGGGCATTGAAGTGTGTGTGCAACCGGGGC 1258

Qy      781 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
      |||
Db      1259 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 1318

Qy      841 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 900
      |||
Db      1319 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1378

Qy      901 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
      |||
Db      1379 CAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 1438

Qy      961 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC 1020
      |||
Db      1439 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCAACTCCCAGAGCAGCAGCTCC 1498

Qy      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
      |||
Db      1499 CCCAGCTAGAGCAGGTCTGCCAGCACCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1558
```

RESULT 2  
AX714139  
LOCUS AX714139 2703 bp DNA linear PAT 15-APR-2003  
DEFINITION Sequence 823 from Patent EP1293569.  
ACCESSION AX714139  
VERSION AX714139.1 GI:29889067  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,  
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,  
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and  
Masuho,Y.  
TITLE Full-length cDNAs  
JOURNAL Patent: EP 1293569-A 823 19-MAR-2003;  
Helix Research Institute (JP) ; Research Association for  
Biotechnology (JP)  
FEATURES Location/Qualifiers  
source 1..2703  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 99.7%; Score 1076.8; DB 2; Length 2703;  
Best Local Similarity 99.8%; Pred. No. 1.3e-309;  
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60  
|||||  
Db 145 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204  
Qy 61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120  
|||||  
Db 205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264  
Qy 121 GAACTCTGGTGGTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGTGTGT 180  
|||||  
Db 265 GAACTCTGGTGGTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGTGTGT 324  
Qy 181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240  
|||||  
Db 325 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384  
Qy 241 CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 300  
|||||  
Db 385 CTGATCGCTTACCGAGAAGCCCAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTG 444  
Qy 301 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 360  
|||||  
Db 445 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACTCCTCCCCCA 504  
Qy 361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCA 420  
|||||  
Db 505 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCA 564  
Qy 421 GGTGGCAGTCCCCCGGCATCGATCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC 480  
|||||  
Db 565 GGTGGCAGTCCCCCGGCATCGATCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC 624  
Qy 481 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 540  
|||||  
Db 625 TTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCAAGCATCGCTGACCCTGATCCCTCT 684  
Qy 541 GACCTACAGTTGACCGAGCAGCCACCAAGCCCGAGGATGGAGCCAGTGGCTCTGTG 600  
|||||  
Db 685 GACCTACAGTTGACCGAGCAGCCACCAAGCCCGAGGATGGAGCCAGTGGCTCTGTG 744  
Qy 601 GCTGGCCTGGGGAGCTGGACCCGGGGGCTTCTGGACAAAGATGCAGAATGTAGGGAG 660  
|||||  
Db 745 GCTGGCCTGGGGAGCTGGACCCGGGGGCTTCTGGACAAAGATGCAGAATGTAGGGAG 804

Qy 661 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720  
 Db 805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864

Qy 721 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 780  
 Db 865 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924

Qy 781 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840  
 Db 925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 984

Qy 841 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 900  
 Db 985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044

Qy 901 CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960  
 Db 1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA 1104

Qy 961 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1020  
 Db 1105 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCC 1164

Qy 1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080  
 Db 1165 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1224

# RESULT 3

AK056285

LOCUS AK056285 2703 bp mRNA linear PRI 20-JAN-2006

DEFINITION Homo sapiens cDNA FLJ31723 fis, clone NT2RI2006682, weakly similar to Human WW domain binding protein-1 mRNA.

ACCESSION AK056285

VERSION AK056285.1 GI:16551642

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS

1  
 Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T.O., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

TITLE

Complete sequencing and characterization of 21,243 full-length human cDNAs

JOURNAL

Nat. Genet. 36 (1), 40-45 (2004)

PUBMED

14702039

REFERENCE 2  
AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,  
Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,  
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,  
Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,  
Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,  
Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,  
Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,  
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
Nagahari,K., Masuho,Y.; Nagai,K. and Isogai,T.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 2703)  
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.  
TITLE Direct Submission  
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:flj-cdna@nifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.  
FEATURES Location/Qualifiers  
source 1. .2703  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RI2006682"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RI2"  
/note="cloning vector: pME18SFL3  
mRNA from NT2 neuronal precursor cells treated 2-weeks  
mitotic inhibitor after 5-weeks retinoic acid (RA)  
induction.  
majorly NT2 neuron"  
ORIGIN  
Query Match 99.7%; Score 1076.8; DB 5; Length 2703;  
Best Local Similarity 99.8%; Pred. No. 1.3e-309;  
Matches 1078; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60  
|||||  
Db 145 ATGCCTTTCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 204  
Qy 61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120  
|||||  
Db 205 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 264  
Qy 121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180  
|||||  
Db 265 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 324  
Qy 181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240  
|||||  
Db 325 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 384  
Qy 241 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTACGGTTTTG 300  
|||||  
Db 385 CTGATCGCTTACCGAGAAGCCACAATTACTCAGCGCTGCCATTTTATTTACGGTTTTG 444  
Qy 301 CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA 360  
|||||  
Db 445 CCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCA 504  
Qy 361 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 420  
|||||  
Db 505 CCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCA 564  
Qy 421 GGTGGCAGTCCCCCGGCATCGATCCCACAGGGGATCCAGGGGGCACAGAGCAGCCCC 480  
|||||

```

Db      565 GGTGGCAGTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCC 624
Qy      481 TTGTCTGAGCCCAGCAGAAGCAGCACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCT 540
      |||
Db      625 TTGTCTGAGCCCAGCAGAAGCAGCACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCT 684
Qy      541 GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 600
      |||
Db      685 GACCTACCAGTTGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTG 744
Qy      601 GCTGGCCTGGGGGAGCTGGACCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 660
      |||
Db      745 GCTGGCCTGGGGGAGCTGGACCGGGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAG 804
Qy      661 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 720
      |||
Db      805 GAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCT 864
Qy      721 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 780
      |||
Db      865 GGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGC 924
Qy      781 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 840
      |||
Db      925 CACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCC 984
Qy      841 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 900
      |||
Db      985 CTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGT 1044
Qy      901 CAGTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCA 960
      |||
Db      1045 CAGCCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCTCA 1104
Qy      961 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1020
      |||
Db      1105 TGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCACTCCCAGAGCAGCAGCTCC 1164
Qy      1021 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
      |||
Db      1165 CCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1224

```

#### RESULT 4

CQ413042

LOCUS CQ413042 3095 bp DNA linear PAT 23-JAN-2004

DEFINITION Sequence 20113 from Patent WO0170979.

ACCESSION CQ413042

VERSION CQ413042.1 GI:41320823

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1

AUTHORS Lee, J. and Lillie, J.

TITLE Genes, compositions, kits, and method for identification,  
assessment, prevention, and therapy of ovarian cancer

JOURNAL Patent: WO 0170979-A 20113 27-SEP-2001;  
Millennium Pharmaceuticals, Inc. (US)

FEATURES

source

Location/Qualifiers

1. .3095

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 97.6%; Score 1054.4; DB 2; Length 3095;

Best Local Similarity 99.9%; Pred. No. 6.5e-303;

Matches 1055; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84

Db 109 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 168

Qy	85	TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG	144
Db	169	TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG	228
Qy	145	TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC	204
Db	229	TGGACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCAC	288
Qy	205	CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC	264
Db	289	CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC	348
Qy	265	AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAAACTATTTACTACCTCCTTAT	324
Db	349	AATTACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCAAACTATTTACTACCTCCTTAT	408
Qy	325	GAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATACAGTGCCTTCAGCTACAG	384
Db	409	GAGGAAGTGGTGAACCGACCTCCAACCTCTCCCCACCATACAGTGCCTTCAGCTACAG	468
Qy	385	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT	444
Db	469	CAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGAT	528
Qy	445	CCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGCTGAGCCAGCAGAAGCAGC	504
Db	529	CCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGCTGAGCCAGCAGAAGCAGC	588
Qy	505	ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC	564
Db	589	ACAAGACCCCCAAGCATCGCTGACCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCC	648
Qy	565	ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCTGGGGAGCTGGACCCG	624
Db	649	ACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCTGGGGAGCTGGACCCG	708
Qy	625	GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT	684
Db	709	GGGGCCTTCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCT	768
Qy	685	GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT	744
Db	769	GAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGT	828
Qy	745	GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG	804
Db	829	GACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAG	888
Qy	805	TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT	864
Db	889	TTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCAT	948
Qy	865	GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTAGTCTCTGAGGAGCAGGCTCGA	924
Db	949	GTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTAGTCTCTGAGGAGCAGGCTCGA	1008
Qy	925	GAGCCTGGGCACCCGACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC	984
Db	1009	GAGCCTGGGCACCCGACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAAC	1068
Qy	985	GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAG	1044
Db	1069	GAGCAGGACTCTCCCAACTCCCAGAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAG	1128
Qy	1045	CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA	1080
Db	1129	CACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA	1164

RESULT 5  
 BD275013  
 LOCUS BD275013 4054 bp DNA linear PAT 17-JUL-2003  
 DEFINITION 50 Human Secreted Proteins.  
 ACCESSION BD275013  
 VERSION BD275013.1 GI:33084781  
 KEYWORDS JP 2002542766-A/13.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 4054)  
 AUTHORS Komatsoulis,G., Rosen,C.A. and Ruben,S.M.  
 TITLE 50 Human Secreted Proteins  
 JOURNAL Patent: JP 2002542766-A 13 17-DEC-2002;  
 Rosen et al

COMMENT OS Homo sapiens  
 PN JP 2002542766-A/13  
 PD 17-DEC-2002  
 PF 09-MAR-2000 JP 2000605624  
 PR 12-MAR-1999 US 60/124093,23-NOV-1999 US 60/166989 PI  
 george komatsoulis,craig a rosen,steven m ruben CC  
 FH Key Location/Qualifiers.

FEATURES  
 source Location/Qualifiers  
 1. .4054  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 96.6%; Score 1043.4; DB 2; Length 4054;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-299;  
 Matches 1055; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 25 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 84  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 107 CAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACAC 166

QY 85 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 144  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 167 TGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTG 226

QY 145 TGGACCATCATCATCATCCTGAGTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 204  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 227 TGGACCATCATCATCATCCTGAGTGCTGCTGTGTTTGCCACCACCGCCGAGCCAAGCAC 286

QY 205 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 264  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 287 CGCCTTCAGGCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCAC 346

QY 265 AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACTATTTACTACCTCC-TTA 323  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 347 AATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACTATTTACTACCTCCTTTA 406

QY 324 TGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACA 383  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 407 TGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACAGTGCCTTCCAGCTACA 466

QY 384 GCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGA 443  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 467 GCAGCAGCAGTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGA 526

QY 444 TCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCCTTGCTCTGAGCCAGCAGAAGCAG 503  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 527 TCCCACAGGGGATCCCAGGGGGCACAGAGCAGCCCCTTGCTCTGAGCCAGCAGAAGCAG 586

QY 504 CACAAGACCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAAGTTGACCGAGCAGC 563  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 587 CACAAGACCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAAGTTGACCGAGCAGC 646

QY 564 CACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCC 623  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 647 CACCAAAGCCCCAGGGATGGAGCCCAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCC 706

QY 624 GGGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTC 683  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 707 GGGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTC 766

QY 684 TGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGG 743  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 767 TGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGG 826

```

Qy      744 TGA CTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGA 803
          |||
Db      827 TGA CTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGA 886

Qy      804 GTTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCA 863
          |||
Db      887 GTTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACTTCTGCGACAGCTGCCA 946

Qy      864 TGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCG 923
          |||
Db      947 TGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTGAGTCTCTGAGGAGCAGGCTCG 1006

Qy      924 AGAGCCTGGGCACCCGACCTGCCACGGCCGCCCGCATGCCTGCTGTGCTGAACACCATCAA 983
          |||
Db     1007 AGAGCCTGGGCACCCGACCTGCCACGGCCGCCCGCATGCCTGCTGTGCTGAACACCATCAA 1066

Qy      984 CGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTTGCCA 1043
          |||
Db     1067 CGAGCAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTTGCCA 1126

Qy     1044 GCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1080
          |||
Db     1127 GCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGGGA 1163

```

RESULT 6

CQ720290

LOCUS CQ720290 1002 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 6224 from Patent WO02068579.

ACCESSION CQ720290

VERSION CQ720290.1 GI:42281147

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

TITLE Kits, such as nucleic acid arrays, comprising a majority of  
human exons or transcripts, for detecting expression and other uses  
thereof

JOURNAL Patent: WO 02068579-A 6224 06-SEP-2002;  
PE Corporation (NY) (US)

FEATURES

source Location/Qualifiers  
1..1002  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 92.8%; Score 1002; DB 2; Length 1002;  
Best Local Similarity 100.0%; Pred. No. 2e-287;  
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      28 GATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACTGC 87
          |||
Db      1 GATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACATCTGTGACACAGGACTGC 60

Qy      88 TGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTGTGG 147
          |||
Db      61 TGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCTGGTGGTTCTGGCTGGTGTGG 120

Qy     148 ACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCACC GC 207
          |||
Db     121 ACCATCATCATCATCCTGAGCTGCTGCTGTGTTGCCACCACCGCCGAGCCAAGCACC GC 180

Qy     208 CTTCAGGCCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCCAAT 267
          |||
Db     181 CTTCAGGCCCCAGCAGCGGCAACATGAAATCAACCTGATCGCTTACCGAGAAGCCCAAT 240

Qy     268 TACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCTAACTATTTACTACCTCCTTATGAG 327
          |||
Db     241 TACTCAGCGCTGCCATTTTATTTAGGTTTTTGCCTAACTATTTACTACCTCCTTATGAG 300

Qy     328 GAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATAACAGTGCCTTCCAGTACAGCAG 387

```



```

Db      301 |||||GAAGTGGTGAACCGACCTCCAACTCCTCCCCACCATAACAGTGCCTTCCAGCTACAGCAG 360
Qy      388 CAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGGCATCGATCCC 447
Db      361 |||||CAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCAGTCCCCCGGGCATCGATCCC 420
Qy      448 ACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGCACA 507
Db      421 |||||ACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGCACA 480
Qy      508 AGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCCACC 567
Db      481 |||||AGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCCACC 540
Qy      568 AAAGCCCCAGGGATGGAGCCCACTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCGGGG 627
Db      541 |||||AAAGCCCCAGGGATGGAGCCCACTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCGGGG 600
Qy      628 GCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCTGAA 687
Db      601 |||||GCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCTGAA 660
Qy      688 CACGGCGCACCCGACAGCAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGTGAC 747
Db      661 |||||CACGGCGCACCCGACAGCAAGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGTGAC 720
Qy      748 TCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTC 807
Db      721 |||||TCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTC 780
Qy      808 AACACACTCATCGATGATGCTCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCATGTG 867
Db      781 |||||AACACACTCATCGATGATGCTCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCATGTG 840
Qy      868 CGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCTAGTCTCTGAGGAGCAGGCTCGAGAG 927
Db      841 |||||CGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCTAGTCTCTGAGGAGCAGGCTCGAGAG 900
Qy      928 CCTGGGCACCCGACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAACGAG 987
Db      901 |||||CCTGGGCACCCGACCTGCCACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAACGAG 960
Qy      988 CAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAG 1029
Db      961 |||||CAGGACTCTCCCAACTCCCAGAGCAGCAGCTCCCCAGCTAG 1002

```

# RESULT 7

AL358790

```

LOCUS      AL358790              131753 bp    DNA        linear    PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-753C18 on chromosome 10 Contains
three novel genes, the CYP17A1 gene for cytochrome P450 17A1, a
profilin 1 (PFN1) pseudogene, the gene for methyltransferase cyt19,
an unactive progesterone receptor 23 kD (TEBP) pseudogene, a
ribosomal protein L22 (RPL22) pseudogene and a CpG island, complete
sequence.
ACCESSION  AL358790
VERSION    AL358790.22  GI:16972912
KEYWORDS   HTG; CpG island; CYP17A1; cyt19; methyltransferase; PFN1; profilin;
RPL22; TEBP.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE  1 (bases 1 to 131753)
AUTHORS    Tromans,A.
TITLE      Direct Submission
JOURNAL    Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
COMMENT    On Nov 16, 2001 this sequence version replaced gi:15131808.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information

```

on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr10>  
 RP11-753C18 is from the library RPCI-11.3 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [vega@sanger.ac.uk](mailto:vega@sanger.ac.uk)  
 -----

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one subclone; and the assembly was confirmed by restriction digest,  
 except on the rare occasion of the clone being a YAC.

FEATURES	Location/Qualifiers
source	1. .131753 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="10" /clone="RP11-753C18" /clone_lib="RPCI-11.3"
misc_feature	1 /note="Clone_left_end: RP11-753C18"
gene	3299. .43314 /gene="C10orf26" /locus_tag="RP11-753C18.10-001"
mRNA	join(3299. .3467,25043. .25145,36943. .37104,39645. .43314) /gene="C10orf26" /locus_tag="RP11-753C18.10-001" /product="chromosome 10 open reading frame 26" /note="match: ESTs: AW006719.1 BF852297.1 BG258775.1 BG720326.1 BQ307980.1 BU537726.1 BU616707.1 BU618493.1 BU784158.1 match: cDNAs: AK000374.1 AK056285.1 BC026369.1 U43563.1"
CDS	join(3441. .3467,25043. .25145,36943. .37104,39645. .40381) /gene="C10orf26" /locus_tag="RP11-753C18.10-001" /standard_name="OTTHUMP00000020381" /note="match: proteins: O95637 P97764 Q8ROW8 Q8WUP5 Q99J20 Q9NX94" /codon_start=1 /product="chromosome 10 open reading frame 26" /protein_id="CAI52496.1" /db_xref="GI:58864791" /db_xref="UniProtKB/TrEMBL:Q5F2G6" /translation="MPFLLGLRQDKEACVGTNNQSYICDTGHCCGQSQCNNYYYELWW FWLWVTIIIIILSCCCVCHHRRAKHRLQAQQRQHEINLIAYREAHNYSALPFYFRFLPN YLLPPYEEVVNRPPTPPPYSAFQLQQQLLPPQCGPAGGSPPGIDPTRGSQGAQSSP LSEPSRSSTRPPSIADPDPSDLVDRAATKAPGMEPSGSVAGLGELDPGAFLDKDAEC REELLKDDSSSEHGAPDSKEKT PGRHRRFTGDSGIEVCVCNRGRHDDDLKEFNTLIDDA LDGPLDFCDSCHVRPPGDEEEGLCQSSEEQAREPGHPLPRPPACLLNTINEQDSPN SQSSSSPS"
misc_feature	28935. .28972 /gene="C10orf26" /locus_tag="RP11-753C18.10-001" /note="Sequence from uni-directional dGTP big dye terminator reads only."
polyA_signal	43292. .43297 /gene="C10orf26" /locus_tag="RP11-753C18.10-001"
polyA_site	43312 /gene="C10orf26" /locus_tag="RP11-753C18.10-001"
polyA_site	43314 /gene="C10orf26" /locus_tag="RP11-753C18.10-001"

polyA\_site 43317  
gene complement(join(57581. .58035,58558. .58661,59561. .59730,60043. .60258,61086. .61172,61835. .62064,62304. .62442,64115. .64583))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-001"

mRNA complement(join(57581. .58035,58558. .58661,59561. .59730,60043. .60258,61086. .61172,61835. .62064,62304. .62442,64115. .64583))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-001"  
/product="cytochrome P450, family 17, subfamily A, polypeptide 1"  
/note="match: ESTs: AI638522.1 AU120851.1 AV701951.1 AV704321.1 AV704538.1 AV705813.1 AV706535.1 AV706722.1 AV707587.1 AV707827.1 AV709994.1 AV710436.1 AV710723.1 AV711490.1 BM559755.1 BM559917.1 match: cDNAs: AF297650.1 BC032843.1 M14564.1 S75277.1"

polyA\_site complement(57581)  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-001"

CDS complement(join(57752. .58035,58558. .58661,59561. .59730,60043. .60258,61086. .61172,61835. .62064,62304. .62442,64115. .64411))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-001"  
/standard\_name="OTTHUMP00000020382"  
/note="match: proteins: P05093 Q95M62 Q9GLD2"  
/codon\_start=1  
/product="cytochrome P450, family 17, subfamily A, polypeptide 1"  
/protein\_id="CAI52498.1"  
/db\_xref="GI:58864793"  
/db\_xref="InterPro:IPR001128"  
/db\_xref="InterPro:IPR002401"  
/translation="MWELVALLLLTILAYLFWPKRRCPGAKYPKSLLSLPLVGSPLPFLP RHGHMHNNFFKLQKKYGP IYSVRMGTKTTIVGHHQLAKEVLIKKGKDFSGRPQMATL DIASNNRKGIAFADSGAHWQLHRRLAMATFALFKDGDQKLEKICQEISTLCDMLATH NGQSIDISFPVFVAVTNVILSICFNSTSYKNGDPELNVIQYNEGIIDNLSKDSLVDLV PWLKIFPNKTLEKLKSHVKIRNDLLNKILENYKEKFRSDSITNMLDTLMQAKMNSDNG NAGPDQDSELLSDNHILTTIGDI FGAGVETTSVVKWTLAFLHNPQVKKKLYEIDQ NVGFSRTPTISDRNRLLEATIREVLRRLRPVAPMLIPHKANVDS SIGEFAVDKGTEV IINLWALHHNEKEWHQPDQFMPERFLNPAGTQLISPSVSYLPFGAGPRSCIGEILARQ ELFLIMAWLLQRFDELPDDGQLPSPLEGIPKVVFLIDSFKVKIKVRQAWREAQAEGST"

gene complement(join(57973. .58035,58558. .58661,59561. .59730,60043. .60206,61086. .61172,61835. .61891))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-003"

mRNA complement(join(57973. .58035,58558. .58661,59561. .59730,60043. .60206,61086. .61172,61835. .61891))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-003"  
/product="cytochrome P450, family 17, subfamily A, polypeptide 1"  
/note="match: ESTs: AV704788.1 AV705037.1 AV706512.1 AV707587.1 AV708270.1"

gene complement(join(57998. .58035,58558. .58753))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-004"

mRNA complement(join(57998. .58035,58558. .58753))  
/gene="CYP17A1"  
/locus\_tag="RP11-753C18.1-004"  
/product="cytochrome P450, family 17, subfamily A, polypeptide 1"  
/note="match: ESTs: BF737959.1"

gene join(<59771. .59921,61227. .61566)  
/locus\_tag="RP11-753C18.3-001"

mRNA join(<59771. .59921,61227. .61566)  
/locus\_tag="RP11-753C18.3-001"  
/product="novel protein"  
/note="match: ESTs: BF529138.1 BG182618.1 BG200234.1 BG201248.1 BG213774.1 BG219080.1 BG220558.1 R16838.1"

CDS join(<59847. .59921,61227. .61466)  
/locus\_tag="RP11-753C18.3-001"  
/standard\_name="OTTHUMP00000020383"

```

/codon_start=1
/product="novel protein"
/protein_id="CAI52497.1"
/db_xref="GI:58864792"
/db_xref="UniProtKB/TrEMBL:Q5F2G5"
/translation="WLWGPASPASRRWGSSTLGSKPTTAWLPESLLEADLSFLSLEVPT
PPLTFSQNGRQDGDTERVYHDDDNQKQRRRAWSPPLLPHSIKHQALTLLCVQPLGVR
ET"
gene      complement(join(60525. .62064,62304. .62643,64115. .64464))
          /gene="CYP17A1"
          /locus_tag="RP11-753C18.1-002"
mRNA      complement(join(60525. .62064,62304. .62643,64115. .64464))
          /gene="CYP17A1"
          /locus_tag="RP11-753C18.1-002"
          /product="cytochrome P450, family 17, subfamily A,
          polypeptide 1"
          /note="match: cDNAs: AK094106.1"
polyA_signal 61531. .61536
polyA_site   /locus_tag="RP11-753C18.3-001"
polyA_site   61558
polyA_site   /locus_tag="RP11-753C18.3-001"
polyA_site   61566

```

```

Query Match      73.1%; Score 789.8; DB 5; Length 131753;
Best Local Similarity 99.7%; Pred. No. 3e-223;
Matches 791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      288 TTTCAGGTTTTTGCCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCC 347
      ||| |||||
Db      39640 TTTAGGGTTTTTGCCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCC 39699
Qy      348 AACTCCTCCCCACCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACA 407

```

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006
---------------------------------

<http://es/ScoreAccessWeb/GetItem.action?AppId=10729895&seqId=1038737&ItemName...> 1/22/2007

3	786	72.8	2702	6	AK078491	AK078491 Mus muscu
4	786	72.8	3871	6	AK142695	AK142695 Mus muscu
5	766.8	71.0	3893	6	AK154686	AK154686 Mus muscu
6	758	70.2	3873	6	AK087487	AK087487 Mus muscu
7	708.2	65.6	735	9	DB293556	DB293556 DB293556
8	683	63.2	800	8	CX163170	CX163170 HESC2_2_F
9	660.8	61.2	3955	6	AK169294	AK169294 Mus muscu
10	587	54.4	587	9	DA190287	DA190287 DA190287
11	578	53.5	712	8	CN427264	CN427264 170005321
12	575	53.2	575	9	DA553446	DA553446 DA553446
13	571.6	52.9	581	3	BP242855	BP242855 BP242855
14	557	51.6	565	3	BP221694	BP221694 BP221694
15	550	50.9	550	9	DA576326	DA576326 DA576326
16	550	50.9	550	9	DB283277	DB283277 DB283277
17	541	50.1	541	9	DB133804	DB133804 DB133804
18	531.6	49.2	852	9	DN935380	DN935380 AGENCOURT
19	526.4	48.7	876	9	DN948780	DN948780 AGENCOURT
20	519.4	48.1	566	9	DA806792	DA806792 DA806792
21	512	47.4	1006	3	BQ054184	BQ054184 AGENCOURT
22	510.4	47.3	680	8	CV868359	CV868359 PDUts1054
23	497.4	46.1	796	5	CF737347	CF737347 UI-M-HD0-
24	496.4	46.0	553	9	DA760493	DA760493 DA760493
25	493.4	45.7	560	9	DA756667	DA756667 DA756667
26	493.4	45.7	762	4	CA316710	CA316710 UI-M-FW0-
27	488	45.2	488	9	DA596718	DA596718 DA596718
28	466.4	43.2	563	9	DA321605	DA321605 DA321605
29	459.2	42.5	844	8	CV118164	CV118164 AGENCOURT
30	454.8	42.1	525	9	CX597272	CX597272 CT020017B
31	450	41.7	450	14	AY405369	AY405369 Homo sapi
32	447.6	41.4	625	4	CA334353	CA334353 NISC_ls08
33	444	41.1	537	9	DB129660	DB129660 DB129660
34	442.4	41.0	837	3	BQ924256	BQ924256 AGENCOURT
35	437	40.5	580	9	DA823881	DA823881 DA823881
36	430.8	39.9	434	4	CB145438	CB145438 K-EST0200
37	428.4	39.7	850	9	DB292857	DB292857 DB292857
38	425	39.4	565	9	DA196239	DA196239 DA196239
39	423.2	39.2	557	9	DA860698	DA860698 DA860698
40	417	38.6	556	9	DB226414	DB226414 DB226414
41	415.2	38.4	551	9	DA887248	DA887248 DA887248
42	414	38.3	553	9	DA723361	DA723361 DA723361
43	410.2	38.0	716	3	BU701616	BU701616 UI-M-F10-
c 44	410	38.0	410	1	AI459545	AI459545 ao86g11.x
45	409.4	37.9	580	9	DN747882	DN747882 GL-Cf-467

#### ALIGNMENTS

RESULT 1  
CR859701

LOCUS CR859701 2792 bp mRNA linear HTC 12-NOV-2004  
 DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp469N2325 (from clone DKFZp469N2325).  
 ACCESSION CR859701  
 VERSION CR859701.1 GI:55730274  
 KEYWORDS HTC.  
 SOURCE Pongo pygmaeus (orangutan)  
 ORGANISM Pongo pygmaeus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Pongo.  
 REFERENCE 1 (bases 1 to 2792)  
 AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
 Fobo, G., Han, M. and Wiemann, S.  
 CONSRTM The German cDNA Consortium  
 TITLE Direct Submission  
 JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
 Neuherberg, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by Agowa (Berlin/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.  
 This clone (DKFZp469N2325) is available at the RZPD Deutsches  
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
 Please contact RZPD for ordering:  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469N2325>  
 Further information about the clone and the sequencing project is

available at <http://mips.gsf.de/projects/cdna/>.

FEATURES  
 source  
 1. .2792  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp469N2325"  
 /tissue\_type="kidney"  
 /clone\_lib="469 (synonym: pkidl). Vector pSport1\_Sfi; host DH10B; sites SfilA + SfilB"  
 /dev\_stage="adult"  
 /note="hypothetical protein (Mus musculus), differentially spliced"  
 gene  
 1. .2792  
 /gene="DKFZp469N2325"  
 CDS  
 43. .360  
 /gene="DKFZp469N2325"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAH91860.1"  
 /db\_xref="GI:55730275"  
 /translation="MERRRLLGGMALLLLQALPSPLSARAEPPLFGGASCGSLSCVW  
 LTTPLLQYWGFGQGRIRKPVWVPTIKATSVTQDTAVDSLAAATTTMNSGGSGWCGPS  
 SSS"

# ORIGIN

Query Match 95.2%; Score 1027.8; DB 6; Length 2792;  
 Best Local Similarity 97.5%; Pred. No. 1.3e-255;  
 Matches 1044; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```

Qy      8 TCCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACA 67
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     202 TACTGGGGGTTTCAGGGCAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAAAGCTACA 261

Qy      68 TCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCT 127
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     262 TCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTATGAACTCT 321

Qy     128 GGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTTGCCACC 187
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     322 GGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTTGCCACC 381

Qy     188 ACCGCCGAGCCAAGCACC GCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCG 247
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     382 ACCGCCGAGCCAAGCACC GCCTTCAGGCCAGCAGCGGCAACATGAAATCAACCTGATCG 441

Qy     248 CTTACCGAGAAGCCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACT 307
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     442 CTTACCGAGAAGCCCACAATTACTCAGCGCTGCCATTTTATTTTCAGGTTTTTGCCAACT 501

Qy     308 ATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACA 367
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     502 ATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAACCTCCTCCCCACCATACA 561

Qy     368 GTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCA 427
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     562 GTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACAGTGTGGCCCTGCAGGTGGCA 621

Qy     428 GTCCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTG 487
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     622 GTCCCCCGGGCGTTGATCCCACCAGGGGATCCCAGGGGGCACAGAGCAGCCCTTGTCTG 681

Qy     488 AGCCCAAGCAGAGCAGCACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTAC 547
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     682 AGCCCAAGCAGAGCAGCACAAGACCCCCAAGCATCGCTGACCCTGATCCCTCTGACCTAC 741

Qy     548 CAGTTGACCGAGCAGCCACCAAGGCCAGGGATGGAGCCCAAGTGGCTCTGTGGCTGGCC 607
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     742 CAGTTGACCGAGCAGCCACCAAGGCCAGGGATGGAGCCCAAGTGGCTCTGTGGCTGGCC 801

Qy     608 TGGGGGAGCTGGACCCGGGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGC 667
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     802 TGGGGGAGCTGGACCCGGGGGCCTTCCTGGACAAAGATGCAGAATGTAGGGAGGAGCTGC 861

Qy     668 TGAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGAC 727
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 862 TGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAAAGAGAAGACGCCTGGGAGAC 921

Qy 728 ATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATG 787  
 |||

Db 922 ATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGTGTGCAACCGGGGCCACCATG 981

Qy 788 ACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGCTCTGGATGGGCCCTGGACT 847  
 |||

Db 982 ACGATGACCTCAAAGAGTTCAACACGCTCATCGATGATGCTCTGGATGGGCCCTGGACT 1041

Qy 848 TCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGATGAGGAGGAAGGCCTCTGTCAGTCCT 907  
 |||

Db 1042 TCTGCGATAGTGGCATGTGCGGCCCCCTGGTGACGAGGAGGAAGGCCTCTGCCAGCCCT 1101

Qy 908 CTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGC 967  
 | |||

Db 1102 CCGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCCACGGCCGCCCGCATGCCTGC 1161

Qy 968 TGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGAGCTCCCCAGCT 1027  
 |||

Db 1162 TGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCAGAGCAGAGCTCCCCAGCT 1221

Qy 1028 AGAGCAGGTCCTGCCAGCACCCAGCAACTTGGCAAAGCAACCAGGGTAGGG 1078  
 |||

Db 1222 AGAGCAGGCCCTGCCAGCACCCAGGAACCTTGGCAAAGCAACCAGGGTAGGG 1272

## RESULT 2

AK052400

LOCUS AK052400 2603 bp mRNA linear HTC 02-SEP-2005

DEFINITION Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched library, clone:D430003C10 product:hypothetical Proline-rich region/Cytochrome c family heme-binding site containing protein, full insert sequence.

ACCESSION AK052400

VERSION AK052400.1 GI:26342634

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

## REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

PUBMED 10349636

## REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

PUBMED 11042159

## REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

PUBMED 11076861

## REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

## REFERENCE 5

AUTHORS The FANTOM Consortium, the RIKEN Genome Exploration Research Group

Phase I and II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs



JOURNAL REFERENCE 6  
 AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group  
 (Genome Network Core Team) and the FANTOM Consortium.  
 TITLE Antisense Transcription in the Mammalian Transcriptome  
 JOURNAL REFERENCE 7  
 AUTHORS The FANTOM Consortium, Riken Genome Exploration Research Group and  
 Genome Science Group (Genome Network Project Core Group).  
 TITLE The Transcriptional Landscape of the Mammalian Genome  
 JOURNAL REFERENCE 8 (bases 1 to 2603)  
 AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
 Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
 Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
 Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
 Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
 Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
 Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
 Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
 Muramatsu,M. and Hayashizaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,  
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)  
 COMMENT cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/.  
 FEATURES  
 source Location/Qualifiers  
 1. 2603  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:D430003C10"  
 /db\_xref="taxon:10090"  
 /clone="D430003C10"  
 /tissue\_type="lung"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="13 days embryo"  
 CDS 143..1189  
 /note="unnamed protein product; hypothetical Proline-rich  
 region/Cytochrome c family heme-binding site containing  
 protein (InterPro|IPR000694 InterPro|IPR000345, evidence:  
 InterPro)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC34974.1"  
 /db\_xref="GI:26342635"  
 /translation="MPFLWGLRQDKEACVGTNNQSYICDTGHCCGQSQCENYYELWW  
 FWLVWTVVILSCCVCHHRRRAKHLQAQQRQHEINLIAYREAHNYSALPFYFRFLPN  
 SLLPPYEYVNRPTPPPPYSAFQLQQQQQLPPPPQGGPPGGSPPGADPPPPQGSQGA  
 QSSPLSGPSRSSTRPPSVADPQSPVPTDREATKASGTESGSPMAGHGELDPGAFLDQ  
 DSECKEELLKDSRERGVSPPSEDKTGPRHRRFTGDSGIEVCVCNRRGHDDDLKEFN  
 TLIDDALDGFLDFCDSCHVRPPVDEEEGLCLSSGQAREHGHPLPRPPACLLNTIN  
 EQDSPNSQHSQSPS"  
 polyA\_signal 2584..2589  
 /note="putative"  
 polyA\_site 2603  
 /note="putative"  
 ORIGIN  
 Query Match 72.8%; Score 786; DB 6; Length 2603;  
 Best Local Similarity 85.0%; Pred. No. 9.6e-193;  
 Matches 933; Conservative 0; Mismatches 145; Indels 20; Gaps 4;

Qy 1 ATGCCTTTCTTTTGGGTCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 60  
 |||  
 Db 143 ATGCCTTTCTCTGGGGCTTAGACAGGATAAGGAAGCCTGTGTGGGTACCAACAATCAA 202

Qy 61 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 120  
 |||  
 Db 203 AGCTACATCTGTGACACAGGACACTGCTGTGGACAGTCTCAGTGCTGCAACTACTACTAT 262

Qy 121 GAACTCTGGTGGTTCTGGCTGGTGTGGACCATCATCATCCTGAGCTGCTGCTGTGTT 180  
 |||  
 Db 263 GAACTCTGGTGGTTCTGGCTGGTGTGGACCGTCGTCATCATTCTGAGCTGCTGCTGTGTC 322

Qy 181 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCCAGCAGCGGCAACATGAAATCAAC 240  
 |||  
 Db 323 TGCCACCACCGCCGAGCCAAGCACCGCCTTCAGGCTCAGCAGCGGCAACATGAAATCAAC 382

Qy 241 CTGATCGCTTACCGAGAAGCCACAACTACTCAGCGCTGCCATTTTATTTCAGGTTTTTG 300  
 |||  
 Db 383 CTGATCGCTTACCGGAAGCCACAACTACTCAGCGTGCCTTTTACTTCAGGTTTTTG 442

Qy 301 CCAAATATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAATCCTCCCCA 360  
 |||  
 Db 443 CCAAATCTTTGCTACCTCCTTATGAGGAAGTGGTGAACCGACCTCCAATCCTCCCCA 502

Qy 361 CCATACAGTGCCTTCAGCTACAGCAGCAGCAGTGCTG-----CCTCCACAGTGT 411  
 |||  
 Db 503 CCGTACAGTGCCTTCAGCTCCAACAGCAGCAGCAGTGCTGCCACCTCCTCCTCAGGT 562

Qy 412 GGCCCTGCAGGTGGCAGTCCCCGGGCATCGA---TCCCACCAGGGGATCCCAGGGGGCA 468  
 |||  
 Db 563 GGCCCTCCAGTGGCAGCCCCCAGGCGCTGACCCGCCGCCAGGGCTCCCAGGGAGCG 622

Qy 469 CAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGCACAAGACCCCCAAGCATCGCTGAC 528  
 |||  
 Db 623 CAGAGCAGTCTTTGTCCGACCGAGCCGAAGCAGCACAAGACCTCCAAGTGTGCGAGAT 682

Qy 529 CCTGATCCCTCTGACCTACCAGTTGACCGAGCAGCCACCAAAGCCCCAGGGATGGAGCCC 588  
 |||  
 Db 683 CCTCAGTCCCTGAAGTGCCCACTGACCGGGAAGCCACCAAAGCCTCTGGGACGGAGTCT 742

Qy 589 AGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCGGGGGCTTCTTGACAAAGATGCA 648  
 |||  
 Db 743 GGTAGCCCATGGCAGGCCACGGGAGCTGGACCCGGGCGCCTTCTTGACACAGGATTCC 802

Qy 649 GAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCTGAACACGGCG-----CACCCGAC 702  
 |||  
 Db 803 GAGTGCAAGGAGGAGCTACTGAAAGATTGCGGTTCTGAGCGTGGCGGCGTGTCCCTGAC 862

Qy 703 AGCAAAGAGAAGACGCTGGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTG 762  
 |||  
 Db 863 AGTGAAGACAAGACCCCGCAGGCATCGCCGCTTCACAGGTGACTCAGGCATTGAGGTG 922

Qy 763 TGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGAT 822  
 |||  
 Db 923 TGTGTGTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTAACACGCTCATAGAT 982

Qy 823 GATGCTCTGGATGGGCCCCTGGACTTCTGCGACAGCTGCCATGTGCGGCCCCCTGGTGAT 882  
 |||  
 Db 983 GATGCTCTGGACGGGCCCCTGGACTTCTGTGACAGCTGTCATGTACGGCCTCCTGTTGAT 1042

Qy 883 GAGGAGGAAGGCTCTGTGAGTCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCAC 942  
 |||  
 Db 1043 GAAGAGGAAGGTCTCTGCCTGTCTCGGAGGTCAGGCTCGGAGCATGGGCACCCCCAC 1102

Qy 943 CTGCCACGGCCGCCGCATGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAAC 1002  
 |||  
 Db 1103 CTGCCACGGCCGCCGCTGCTGCTGCTAAACACCATCAATGAGCAGGACTCTCCAAAT 1162

Qy 1003 TCCCAGAGCAGCAGCTCCCCAGCTAGAGCAGGTCTGCCAGCACCAGCAACTTGGCAA 1062  
 |||  
 Db 1163 TCCCAGACAGTGGCTCCCCGAGCTAGGACACGCCCTGCCTGCACCTAGGAGCTCAACGA 1222

Qy 1063 AGCAACCAGGGTAGGGGA 1080  
 |||  
 Db 1223 --CAACCAGGGTCAGGGA 1238

RESULT 3  
AK078491  
LOCUS AK078491 2702 bp mRNA linear HTC 02-SEP-2005  
DEFINITION Mus musculus 12 days embryo female mullerian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6820415N12 product:hypothetical Proline-rich region/Cytochrome c family heme-binding site containing protein, full insert sequence.  
ACCESSION AK078491  
VERSION AK078491.1 GI:26347312  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
PUBMED 11076861  
REFERENCE 4  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE 5  
AUTHORS The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
REFERENCE 6  
AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.  
TITLE Antisense Transcription in the Mammalian Transcriptome  
JOURNAL Science 309, 1564-1566 (2005)  
REFERENCE 7  
AUTHORS The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).  
TITLE The Transcriptional Landscape of the Mammalian Genome  
JOURNAL Science 309, 1559-1563 (2005)  
REFERENCE 8 (bases 1 to 2702)  
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.





PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
PUBMED 11076861  
REFERENCE 4  
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.

CONSRM RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)  
PUBMED 11217851  
REFERENCE 5  
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

CONSRM FANTOM Consortium

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420 (6915), 563-573 (2002)  
PUBMED 12466851  
REFERENCE 6  
AUTHORS Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzius,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavolan,M., Davis,M.J., Wilming,L.G., Aidinis,V., Allen,J.E.,

Ambesi-Impiombato,A., Apweiler,R., Aturaliya,R.N., Bailey,T.L., Bansal,M., Baxter,L., Beisel,K.W., Bersano,T., Bono,H., Chalk,A.M., Chiu,K.P., Choudhary,V., Christoffels,A., Clutterbuck,D.R., Crowe,M.L., Dalla,E., Dalrymple,B.P., de Bono,B., Della Gatta,G., di Bernardo,D., Down,T., Engstrom,P., Fagiolini,M., Faulkner,G., Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E., Gustincich,S., Harbers,M., Hayashi,Y., Hensch,T.K., Hirokawa,N., Hill,D., Huminiecki,L., Iacono,M., Ikeo,K., Iwama,A., Ishikawa,T., Jakt,M., Kanapin,A., Katoh,M., Kawasaki,Y., Kelso,J., Kitamura,H., Kitano,H., Kollias,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K., Kurochkin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liuni,S., McWilliam,S., Madan Babu,M., Madera,M., Marchionni,L., Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Mottagui-Tabar,S., Mulder,N., Nakano,N., Nakauchi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J., Pavesi,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Rost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C., Schonbach,C., Sekiguchi,K., Semple,C.A., Seno,S., Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sinclair,B., Sperling,S., Stupka,E., Sugiura,K., Sultana,R., Takenaka,Y., Taki,K., Tammoja,K., Tan,S.L., Tang,S., Taylor,M.S., Tegner,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,E., Verardo,R., Wei,C.L., Yagi,K., Yamanishi,H., Zabarovsky,E., Zhu,S., Zimmer,A., Hide,W., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T., Brusic,V., Quackenbush,J., Wahlestedt,C., Mattick,J.S., Hume,D.A., Kai,C., Sasaki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M., Suzuki,M., Aoki,J., Arakawa,T., Iida,J., Imamura,K., Itoh,M., Kato,T., Kawaji,H., Kawagashira,N., Kawashima,T., Kojima,M., Kondo,S., Konno,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessy,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watahiki,A., Okamura-Oho,Y., Suzuki,H., Kawai,J. and Hayashizaki,Y.

CONSRTM FANTOM Consortium  
TITLE The transcriptional landscape of the mammalian genome  
JOURNAL Science 309 (5740), 1559-1563 (2005)  
PUBMED 16141072  
REFERENCE 7  
AUTHORS Katayama,S., Tomaru,Y., Kasukawa,T., Waki,K., Nakanishi,M., Nakamura,M., Nishida,H., Yap,C.C., Suzuki,M., Kawai,J., Suzuki,H., Carninci,P., Hayashizaki,Y., Wells,C., Frith,M., Ravasi,T., Pang,K.C., Hallinan,J., Mattick,J., Hume,D.A., Lipovich,L., Batalov,S., Engstrom,P.G., Mizuno,Y., Faghihi,M.A., Sandelin,A., Chalk,A.M., Mottagui-Tabar,S., Liang,Z., Lenhard,B. and Wahlestedt,C.

CONSRTM RIKEN Genome Exploration Research Group  
TITLE Antisense transcription in the mammalian transcriptome  
JOURNAL Science, 309 (5740), 1564-1566 (2005)  
PUBMED 16141073  
REFERENCE 8 (bases 1 to 3871)  
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hori,F., Iida,J., Imamura,K., Imotani,K., Itoh,M., Kanagawa,S., Kawai,J., Kojima,M., Konno,H., Murata,M., Nakamura,M., Ninomiya,N., Nishiyori,H., Nomura,K., Ohno,M., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shiraki,T., Tagami,M., Tagami,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission  
JOURNAL Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/.

FEATURES  
source Location/Qualifiers  
1..3871  
/organism="Mus musculus"  
/mol\_type="mRNA"





[illegible]

## RESULT 5

AK154686

LOCUS	AK154686	3893 bp	mRNA	linear	HTC 21-SEP-2005
-------	----------	---------	------	--------	-----------------

DEFINITION Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F630103G18 product:hypothetical protein, full insert sequence.

ACCESSION AK154686

VERSION AK154686.1 GI:74185772

KEYWORDS. HTC; CAP trapper.

SOURCE *Mus musculus* (house mouse)

ORGANISM	Mus musculus
----------	--------------

Eumarkyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

## REFERENCE

Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

PUBMED 10349636

## REFERENCE

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
-------	------------------------------------------------------------------------------------------------------------------------------------

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

PUBMED 11042159

## REFERENCE

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

PUBMED	11076861
--------	----------

## REFERENCE

AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Heischmann, W., Gaudes, R., Gissi, S., Ring, B., Roehrig, H.,

Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,  
 Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,  
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
 Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
 Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,  
 Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,  
 Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
 Marchionni,L., Mashima,J., Mazzairelli,J., Mombaerts,P., Nordone,P.,  
 Ring,B., Ringwald,M., Rodríguez,I., Sakamoto,N., Sasaki,H.,  
 Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
 Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,  
 Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.  
 and Hayashizaki,Y.

CONSRTM RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409 (6821), 685-690 (2001)  
 PUBMED 11217851  
 REFERENCE 5  
 AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,  
 Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,  
 Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,  
 Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,  
 Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,  
 Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,  
 Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,  
 Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,  
 Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,  
 Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,  
 Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,  
 Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,  
 Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,

[start](#) | [next page](#)

SCORE 1.3 BuildDate: 11/17/2006

# SCORE Search Results Details for Application 10729895 and Search Result 20070109\_162406\_us-10-729-895a-1.rnpbn.

[Score Home](#)    [Retrieve Application](#)    [SCORE System](#)    [SCORE](#)    [Comments /](#)  
[Page](#)    [List](#)    [Overview](#)    [FAQ](#)    [Suggestions](#)

This page gives you Search Results detail for the Application 10729895 and Search Result 20070109\_162406\_us-10-729-895a-1.rnpbn.

[start](#) | [next page](#)

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2007 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2007, 02:13:47 ; Search time 380 Seconds  
(without alignments)  
7745.241 Million cell updates/sec

Title: US-10-729-895A-1  
Perfect score: 1080  
Sequence: 1 atgcctttccttttgggtct.....aaagcaaccagggtaggga 1080

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3650718 seqs, 1362588608 residues

Total number of hits satisfying chosen parameters: 7301436

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA\_New:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq1:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq2:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq3:\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	789.8	73.1	39315	7	US-11-033-056A-36264	Sequence 36264, A
c 2	430.2	39.8	35944	8	US-11-266-748A-59993	Sequence 59993, A
c 3	430.2	39.8	341983	7	US-11-033-056A-37870	Sequence 37870, A
c 4	430.2	39.8	684973	8	US-11-266-748A-32013	Sequence 32013, A
c 5	366.6	33.9	420	8	US-11-266-748A-350361	Sequence 350361,

6	366.6	33.9	420	8	US-11-266-748A-433740	Sequence 433740,
7	350.4	32.4	352	8	US-11-266-748A-33938	Sequence 33938, A
8	330.4	30.6	435	6	US-10-513-369-5134	Sequence 5134, Ap
9	224	20.7	138251	6	US-10-541-749-96	Sequence 96, Appl
10	123.2	11.4	736	8	US-11-266-748A-193810	Sequence 193810,
11	123.2	11.4	736	8	US-11-266-748A-227330	Sequence 227330,
12	123.2	11.4	810	7	US-11-371-354-58590	Sequence 58590, A
13	123.2	11.4	1000	8	US-11-266-748A-116884	Sequence 116884,
c 14	123.2	11.4	1000	8	US-11-266-748A-159048	Sequence 159048,
c 15	123.2	11.4	1166	8	US-11-266-748A-73679	Sequence 73679, A
c 16	123.2	11.4	1166	8	US-11-266-748A-107945	Sequence 107945,
17	123.2	11.4	1166	8	US-11-266-748A-126490	Sequence 126490,
18	123.2	11.4	2047	8	US-11-293-697-107	Sequence 107, App
19	80.4	7.4	861	7	US-11-478-144-379	Sequence 379, App
20	63	5.8	63	7	US-11-511-035-48751	Sequence 48751, A
21	48	4.4	510	8	US-11-266-748A-73680	Sequence 73680, A
22	48	4.4	510	8	US-11-266-748A-107946	Sequence 107946,
c 23	48	4.4	510	8	US-11-266-748A-126491	Sequence 126491,
24	47	4.4	64	7	US-11-511-035-532998	Sequence 532998,
25	47	4.4	894	8	US-11-266-748A-175387	Sequence 175387,
26	47	4.4	1162	8	US-11-266-748A-247578	Sequence 247578,
27	43.2	4.0	64196	7	US-11-033-056A-38372	Sequence 38372, A
28	40.2	3.7	1000	8	US-11-266-748A-198617	Sequence 198617,
29	40.2	3.7	4681	9	US-11-218-305-1801	Sequence 1801, Ap
30	39.8	3.7	38630	6	US-10-669-920-206	Sequence 206, App
c 31	39.2	3.6	3222	9	US-11-218-305-240	Sequence 240, App
32	38.8	3.6	391	8	US-11-266-748A-179520	Sequence 179520,
33	38.8	3.6	775	6	US-10-374-780A-1099	Sequence 1099, Ap
34	38.8	3.6	775	6	US-10-374-780A-1102	Sequence 1102, Ap
35	38.8	3.6	978	8	US-11-266-748A-352617	Sequence 352617,
c 36	38.8	3.6	978	8	US-11-266-748A-435996	Sequence 435996,
37	38.8	3.6	1000	8	US-11-266-748A-392185	Sequence 392185,
c 38	38.8	3.6	1000	8	US-11-266-748A-482903	Sequence 482903,
39	38.8	3.6	1205	6	US-10-374-780A-1101	Sequence 1101, Ap
40	38.8	3.6	1205	6	US-10-374-780A-1104	Sequence 1104, Ap
41	38.8	3.6	1253	9	US-11-218-305-8594	Sequence 8594, Ap
42	38.8	3.6	2651	8	US-11-266-748A-57574	Sequence 57574, A
c 43	38.6	3.6	1979	6	US-10-612-783-3041	Sequence 3041, Ap
c 44	37.8	3.5	1439	6	US-10-449-902-6346	Sequence 6346, Ap
45	37.8	3.5	3155	8	US-11-266-748A-350021	Sequence 350021,

#### ALIGNMENTS

##### RESULT 1

US-11-033-056A-36264  
 ; Sequence 36264, Application US/11033056A  
 ; Publication No. US20060292572A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: STUART, ROBERT O.  
 ; APPLICANT: STUART, ELIZABETH DUFF  
 ; APPLICANT: WACHSMAN, WILLIAM  
 ; APPLICANT: MERCOLA, DANIEL  
 ; APPLICANT: MCCLELLAND, MICHAEL  
 ; APPLICANT: WANG-RODRIGUEZ, JESSICA  
 ; APPLICANT: TARIN, DAVID  
 ; APPLICANT: BERRY, CHARLES C.  
 ; APPLICANT: ARDEN, KAREN  
 ; APPLICANT: WASSERMAN, LINDA  
 ; APPLICANT: GOODISON, STEVEN  
 ; APPLICANT: KLACANSKY, IGOR  
 ; TITLE OF INVENTION: CELL-TYPE-SPECIFIC PATTERNS OF GENE EXPRESSION  
 ; FILE REFERENCE: 15670-073001  
 ; CURRENT APPLICATION NUMBER: US/11/033,056A  
 ; CURRENT FILING DATE: 2005-01-10  
 ; PRIOR APPLICATION NUMBER: 60/535,382  
 ; PRIOR FILING DATE: 2004-01-09  
 ; PRIOR APPLICATION NUMBER: 60/536,163  
 ; PRIOR FILING DATE: 2004-01-12  
 ; NUMBER OF SEQ ID.NOS: 38888  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 36264  
 ; LENGTH: 39315  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens

Query Match		73.1%;	Score 789.8;	DB 7;	Length 39315;
Best Local Similarity		99.7%;	Pred. No. 2.5e-207;		
Matches	791;	Conservative	0;	Mismatches	2;
				Indels	0;
					Gaps
					0;
Qy	288	TTTCAGGTTTTTGGCCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCC	347		
Db	36342	TTTAGGGTTTTTGGCCAAACTATTTACTACCTCCTTATGAGGAAGTGGTGAACCGACCTCC	36401		
Qy	348	AACTCCTCCCCACCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACA	407		
Db	36402	AACTCCTCCCCACCATACAGTGCCTTCCAGCTACAGCAGCAGCAGCTGCTGCCTCCACA	36461		
Qy	408	GTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGC	467		
Db	36462	GTGTGGCCCTGCAGGTGGCAGTCCCCGGGCATCGATCCCACCAGGGGATCCCAGGGGGC	36521		
Qy	468	ACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGCACAAAGACCCCCAAGCATCGCTGA	527		
Db	36522	ACAGAGCAGCCCTTGTCTGAGCCAGCAGAAGCAGCACAAAGACCCCCAAGCATCGCTGA	36581		
Qy	528	CCCTGATCCCTCTGACCTACCAAGTGGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCC	587		
Db	36582	CCCTGATCCCTCTGACCTACCAAGTGGACCGAGCAGCCACCAAGCCCCAGGGATGGAGCC	36641		
Qy	588	CAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCCTGGACAAAGATGC	647		
Db	36642	CAGTGGCTCTGTGGCTGGCCTGGGGGAGCTGGACCCGGGGGCCTTCCTGGACAAAGATGC	36701		
Qy	648	AGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAA	707		
Db	36702	AGAATGTAGGGAGGAGCTGCTGAAAGATGACAGCTCTGAACACGGCGCACCCGACAGCAA	36761		
Qy	708	AGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGT	767		
Db	36762	AGAGAAGACGCCTGGGAGACATCGCCGCTTCACAGGTGACTCGGGCATTGAAGTGTGTGT	36821		
Qy	768	GTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGC	827		
Db	36822	GTGCAACCGGGGCCACCATGACGATGACCTCAAAGAGTTCAACACACTCATCGATGATGC	36881		
Qy	828	TCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCATGTGCGGGCCCCCTGGTGATGAGGA	887		
Db	36882	TCTGGATGGGCCCCCTGGACTTCTGCGACAGCTGCCATGTGCGGGCCCCCTGGTGATGAGGA	36941		
Qy	888	GGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCC	947		
Db	36942	GGAAGGCCTCTGTCACTCCTCTGAGGAGCAGGCTCGAGAGCCTGGGCACCCGCACCTGCC	37001		
Qy	948	ACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCA	1007		
Db	37002	ACGGCCGCCCGCATGCCTGCTGCTGAACACCATCAACGAGCAGGACTCTCCCAACTCCCA	37061		
Qy	1008	GAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAA	1067		
Db	37062	GAGCAGCAGTCCCCAGCTAGAGCAGGTCTGCCAGCACCCAGCAACTTGGCAAAGCAA	37121		
Qy	1068	CCAGGGTAGGGGA	1080		
Db	37122	CCAGGGTAGGGGA	37134		

US-11-266-748A-59993/c  
; Sequence 59993, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03



